

Análisis Fonoaudiológico para la detección de alzheimer con IA

Juan David Guarnizo Gutierrez, *Ingeniería de Sistemas y Computacion*

Resumen—En Colombia el Sistema de Salud y en general la Academia Colombiana aun tiene mucho por implementar e investigar para beneficiar a la poblacion del pais. En el territorio colombiano los que mas necesidades sociales y medicas poseen son la poblacion de tercera edad, sufran debido a la necesidad de mas investigacion en medicina y la escaza implementacion de nuevas tecnologias, debido a los altos costos y la falta de recursos. Esta falta de calidad hace que los costos del buen vivir sean asequibles para muy pocas personas y en un país en desarrollo como este y en general, en Latinoamerica, la salud de muchos pacientes se deteriora en circunstancias donde esto es evitable. Una de las enfermedades que más afecta a la tercera edad es el Alzheimer una enfermedad neurodegenerativa que puede culminar con la dependencia total del individuo de sus familiares y una muerte con la pérdida de lo que nos hace esencialmente humanos, nuestra mente. La Inteligencia Artificial implica retos importantes dentro de la sociedad colombiana, mientras que a la ves nos brinda muchas soluciones a estos problemas, debido a que con la correcta investigación de la misma y de diferentes modelos de inteligencia artificial, es posible abaratar los costos del diagnóstico del alzheimer y por lo tanto la manutención de personas con esta enfermedad, mejorando con el pasar del tiempo la calidad de vida de muchas personas, en este documento se plantea estudiar los avances en inteligencia artificial en los últimos años, su implementación en la medicina, y el desarrollo de un proyecto que permita ayudar a identificar el alzheimer en pacientes a través de reconocimiento de su voz.

Abstract—In Colombia the Health System and in general the Colombian Academy still has to much to implement and investigate to benefit the population of the country. In the Colombian territory those who have more social and medical needs are the elderly population, they suffer due to the need for more research in medicine and the scarce implementation of new technologies, due to high costs and lack of resources. This lack of quality makes the costs of good living affordable for very few people and in a developing country like this one and in general, in Latin America, the health of many patients deteriorates in circumstances where this is avoidable. One of the diseases that most affects the elderly is Alzheimer's, a neurodegenerative disease that can culminate in the total dependence of the individual on his or her relatives and a death with the loss of what makes us essentially human, our mind. Artificial Intelligence implies important challenges within the Colombian society, while at the same time it offers many solutions to these problems, due to the fact that with the correct investigation of the same and of different models of artificial intelligence, it is possible to reduce the costs of the diagnosis of Alzheimer's and therefore the maintenance of people with this disease, improving with the passing of time the quality of life of many people, in this document we propose to study the advances in artificial intelligence in the last years, and to study the advances in artificial intelligence in the last years.

Translated with DeepL.com (free version)

Index Terms—Alzheimer, dementia, AI, diagnosis , medicine



1. INTRODUCCIÓN

Que es el Alzheimer?

El alzheimer es una enfermedad de alto riesgo que provoca en el paciente perdida de memoria, depresion, falta de capacidad para administrar los recursos, afasia, agnosia, apraxia y, resumiendo, el deterioro general de las capacidades cognitivas, de lenguaje y sociales lo que genera que las personas con la enfermedad dependan de sus seres queridos en esencialmente todo generando altos niveles de stres en ellas mismas y sus familiares. En el pais tenemos nuevos casos cada 4 minutos, siendo un padecimiento con casos mas recurrentes y entre los factores de riesgo podemos hallar alto o poco consumo de alcohol, consumo de cigarrillo y entre otros, depresion o factores geneticos [1] factores de riesgo presentes en la mayoria de la poblacion de

tercera edad.

Entre los retos a los que nos enfrentamos con el Alzheimer, uno de los mas criticos es el diagnostico debido a que los medios actuales de diagnostico suelen ser invasivos y costosos y a que ya con el diagnostico, el precio economico del tratamiento es muy alto. Generalmente el umbral de diagnostico se encuentra apartir de los 60 años y esto es un dato primordial para entender por que, el uso de la tecnologia y el avance de la medicina es esencial para cuidar y proteger la vida de miles de personas de la 3ra edad y sobretodo en paises de tercer mundo debido a que la reduccion de costos pueden beneficiar a la cantidad de personas que pueden acceder a una medicina justa y eficaz, puesto que como se mencionaba anteriormente, los tratamientos son costosos al involucrar el uso de farmacos, antioxidantes y antipsicoticos,

lo cual hace que los costos aumenten conforme progresa la enfermedad y la ausencia de acceso a ellos atacan la calidad de vida y los derechos fundamentales de los pacientes.

Segun un estudio en Estados Unidos [2], el solo valor de diagnostico de la enfermedad en el 91 ya presentaba una serie de costos importantes, y el tratamiento de la enfermedad presenta a un mas costos, no obstante, esto sin considerar los aumentos en inflacion, cuanto antes se logre detectar la enfermedad, sera mucho mas sencillo evitar que la persona se deteriore lo cual mejoraria mucho mas la calidad de vida de la persona y tambien facilitaria la manutencion del paciente al poseer mas control sobre su enfermedad, ademas de brindarle a la persona el beneficio de decidir como va a enfrentarse a la enfermedad cuando aun posee las capacidades mentales para lograr hacerlo.

Annual (undiscounted) direct costs	
Diagnosis (first year only)	\$ 1 450
Nursing home	7 570
Long-term mental hospital	392
Paid home care	3 140
Regular physician care	233
Acute care hospitalization	1 202
Other patient direct costs	0
Caregiver medical care	153
Total direct cost (first year only)	14 140
Total direct cost (second and later years)	12 690
Annual (undiscounted) indirect costs of unpaid home care	
	20 900
Total cost first year, excluding morbidity and mortality	35 040
Total cost second and later years, excluding morbidity and mortality	33 590
Total discounted direct cost*	47 581
Total discounted (direct cost + unpaid caregiver cost)*	123 556
Total discounted (direct cost + unpaid caregiver cost + disability and premature mortality cost)*	173 932

*Assuming a 4% annual discount rate and survival of 3.3 years for men, 4.3 years for women.

Figura 1. Costos de tener Alzheimer

La medicina y la tecnologia

la IA ha logrado avanzar mucho en los ultimos años gracias a los avances en hardware y desarrollo de modelos abstractos, hemos logrado generar modelos que permiten reconocer cada vez mas enfermedades en el cuerpo humano a traves de vision computacional y analisis de imagen [3], con procesamiento de voz o incluso con analisis multimodal [4], esto debido a que en realidad muchas enfermedades ya se diagnostican analizando este tipo de informacion y a que el cuerpo da muchas señales ante la presencia de una enfermedad. En enfermedades neurodegenerativas como el Alzheimer el uso de estas tecnologias es util ya que sus afectaciones se pueden reconocer a traves del habla y procesamiento de lenguaje del paciente y el lenguaje es uno de los indicativos mas importantes del estado social, emocional y cognitivo de una persona [14]. Lo anterior debido a que la velocidad de procesamiento, el lexico y en general la comunicacion ayuda a reconocer patrones patologicos ya que para un uso adecuado del lenguaje es necesario un cerebro sano y funcional[5]. Asi mismo, el analisis de la voz y los sonidos del cuerpo nos brindan varias pistas patologicas en medicina, con el estudio de esto en el pasado se han detectado enfermedades como el cancer de laringe, la estenosis subglotica o torticollis espasmódica [6]. Para lograr la implementacion tecnologica dentro del diagnostico es necesario hallar modelos que permitan encontrar los patrones adecuados y biomarcadores en el lenguaje para distinguir cada caso de alzheimer y ayudar a dar diagnosticos

mas precisos en todos los rangos de avance de la enfermedad, este tipo de modelamiento puede ser realizado mas eficazmente gracias a los ultimos desarrollos de deep learning y supervised learning [7], los cuales permiten estudiar y crear modelos con mayor focalizacion e importancia, lo cual en contra posicion con el MRI, analisis psicologicos y otros exámenes tradicionales es bastante economico y poco invasivo una vez solidificado el modelo.

Todo este avance se ha dado gracias a distintos grupos de investigacion y empresas de Inteligencia Artificial y software que han generado herramientas y repositorios como Hugging face, data2vec [8], bases de datos de Audio y videos de pacientes [13] de facil de acceso y gracias a la implementacion de otras herramientas mas tradicionales como modelos de regresion lineal y algoritmos de optimizacion para optimizar estos modelos [9]. Todo este avance comunitario nos permitira con el tiempo perfeccionar estos sistemas predictivos y lograr cada vez mas avances dentro de la medicina y la ingenieria aplicada a la misma, no obstante, como todo problema medico y de machine learning es necesario una gran cantidad de datos, por lo cual un mayor indice de investigacion e interes por parte de la comunidad cientifico en estas herramientas puede generar que se obtenga una mayor cantidad de datos y se financien mas proyectos para ayudar al ser humano, entrenar modelos no es sencillo y es necesario datos de alta calidad y tener en consideracion otras enfermedades que pueden pasar como falsos positivos de no ser considerados, todo esto mientras se tiene en claro tambien la importancia de la etica y del factor humano dentro de la creacion de estos modelos para evitar sesgos en las predicciones y tambien asegurar la privacidad de los datos del paciente y entender los deseos del mismo adecuadamente, aplicando modelos de inteligencia artificial junto a humanos para asegurar un tratamiento digno y humano del paciente[10], puesto que una falta del conexion con el sector medico, los doctores y los pacientes, estos modelos pueden avanzar mucho pero nunca ser implementados [11].

2. ESTADO DEL ARTE

El diagnostico en alzheimer ha sido investigado principalmente en las ultimas decadas debido al avance tecnologico, una mayor comprension de la enfermedad y a el envejecimiento prolongado de la mayoria de la poblacion, lo que ha incentivado al desarrollo de mas metodos para ayudar a la poblacion con Alzheimer. Convencionalmente se ha usado el MMSE y el MoCa [4][15] los cuales ayudan a diagnosticar a traves de exámenes cognitivos si el paciente posee o no alzheimer, estos exámenes nos pueden ayudar a extraer informacion acerca de como y en que circunstancias podemos obtener datos para determinar el diagnostico del paciente, tambien es posible recurrir a escaneos (MRI) pero estos son mucho mas invasivos y costosos para los pacientes, lo cual incentivo a la comunidad medica a empezar a utilizar otros medios para reconocer enfermedades.

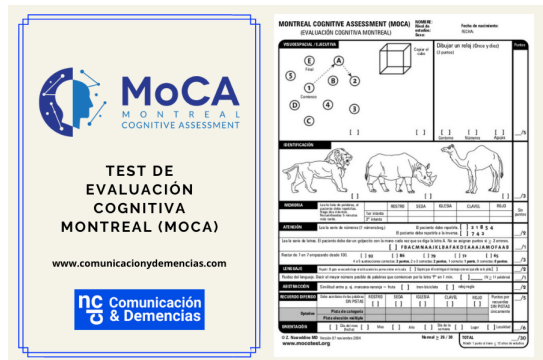


Figura 2. MoCA

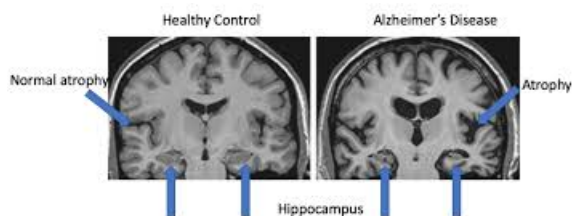


Figura 3. MRI Alzheimer

Otro aspecto de investigación a considerar son los biomarcadores del cuerpo [3][12] los cuales han permitido recolectar mas informacion acerca de como el cuerpo se comporta, como obtener datos de el y obtener informacion valiosa de los mismos. Con el desarrollo tecnologico y la mayor capacidad de procesamiento que nace de la era de la informacion, es muy comun en la actualidad implementar modelos estadísticos, como los modelos de probit y los modelos de regresion. Con la acumulacion de datos valiosos tambien se le empezo a dar mas valor a las redes neuronales, en particular a las redes neuronales recurrentes y convolucionales, ya que pueden complementar facilmente los sistemas de diagnostico tradicionales y tambien nos permite extraer y extrapolar mas informacion que no se habia observado antes por los ojos humanos lo cual ayuda a diagnosticar la enfermedad, investigarla y desarrollar farmacos, los cuales han demostrado ser efectivos para la prevencion y proteccion de los pacientes[1] lo cual demuestra el valor de estos ultimos avances.

En lo que respecta particularmente a la Inteligencia artificial, los transformers y el desarrollo de los modelos con lenguaje natural, son los que mas facilidades han tenido para incorporarse en el campo de la medicina debido a nuestra experiencia con tecnicas como la auscultacion y con el procesamiento de big data dandonos mucho con lo que trabajar[7] ya que la inteligencia artificial multimodal permite procesar todo tipo de datos, tanto imagenes, como audio y texto, lo cual es ideal para combinar diferentes test que se realizan para diagnosticar el alzheimer y asi tambien poder estudiar correlaciones biologicas que han pasado desapercibidas hasta ahora

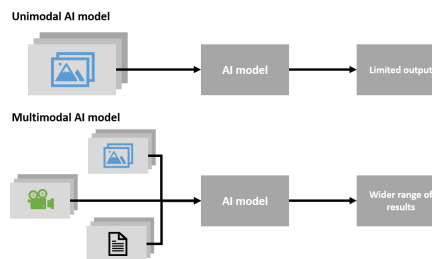


Figura 4. IA multimodal

Para observar mas a profundidad a continuacion estudiaremos brevemente 5 casos y las herramientas que utilizaron para lograr realizar sus respectivas investigaciones

1) Use of Deep Neural Networks to Predict Obesity With Short Audio Recordings: Development and Usability Study El objetivo del estudio fue poder predecir si una persona poseia o no obesidad con la ayuda de un audio, el estudio se postulo de la siguiente manera:

- **Consideraciones eticas:** Se presento el estudio al comite etico de la universidad de Shangai y se obtuvo en consentimiento de cada uno de los participantes
- **Recoleccion de Datos:** Se recolectaron los datos con un encuesta web donde se le solicitaban a los participantes informacion y se les solicitaba enviar audios, se obtuvieron un total de 696 participantes, entre los cuales 500 eran hombres con una edad promedio de 24 años y aproximadamente 400 participantes no poseian obesidad
- **Procesamiento de los datos:** Los audios fueron segmentados y transformados en espectrogramas, se porcionaron los datos en 592 espectrogramas para el entrenamiento y 105 para la validacion de los datos, finalmente se utilizo un 5-fold cross-validation durante el entrenamiento
- **Modelo:** Se adapto el Framework YOLO especializado en procesamiento de imagen, utilizando batch normalization, learning rate optimization y early stopping para optimizar y mejorar el rendimiento de la IA, en comparacion tambien se utilizo una red neuronal convolucional entrenada con datos recolectados con MFCC (un metodo para procesamiento de audio que extrae informacion acustica y armonica importante)
- **Conclusiones:** El modelo logro una precision del 70 % a nivel global, no obstante obtuvo una mayor precision para detectar patrones de no obesidad, por lo cual el modelo necesita explorar mas alternativas de recoleccion de datos, probar con otras tecnicas de procesamiento y estudiar mas a profundidad el tema para lograr aplicar con mayor eficacia las redes neuronales al estudio de la obesidad y la prevencion y diagnostico de la misma

2) Technology COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings Durante la pandemia del Covid y la necesidad de diagnosticar rapidamente a las personas ante la gravedad de la situacion, se realizo un modelo buscando explorar las posibilidades de diagnosticar covid con informacion auditiva recolectada a traves de audios

- **Consideraciones eticas:** El estudio fue realizado considerando los parametros eticos de la IEEE y en especifico la EMB

- **Recoleccion de Datos:** Se creo un sitio web donde se recolecto informacion y muestras acerca de personas con Covid y sin covid con un total de 5230 muestras, con la mitad siendo pacientes positivos para covid y la otra mitad negativos, teniendo un set balanceado en todos los parametros
- **Procesamiento de los datos:** Para el procesamiento de los datos se dividió el audio en 6 segmentos y se utilizó el MFCC para lograr obtener informacion valiosa acerca de los datos, para posteriormente procesar aun mas los datos con una mascara de Poisson, se utilizó el 80% de los datos para el entrenamiento y el 20% para la validacion del modelo
- **Modelo:** El modelo se baso en un esquema de procesamiento por partes donde se utilizaron redes neuronales ResNet50, se utilizaron 4 redes, 3 para procesar y caracterizar diferentes características del audio (la salud de las cuerdas vocales, el tracto respiratorio y los sentimientos) pasar los resultados de esas redes neuronales a una capa de pooling sobre las cuales finalmente, pasaran sus resultados a una ultima shallow network densa que finalmente daría el resultado, se utilizaron datos tambien de Librispeech para pre entrenar las redes para procesamiento de audio y posteriormente se terminaron de entrenar con los datos recolectados
- **Conclusiones:** El modelo fue un exito ya que alcanzo una precision del 98.5 %, evitando los falsos positivos, lo cual puede ayudar a diagnosticar la enfermedad facilmente y sin un costo elevado, demostrando el potencial de la inteligencia artificial para el diagnostico de enfermedades

3) On the Selection of Non-Invasive Methods Based on Speech Analysis Oriented to Automatic Alzheimer Disease Diagnosis

El diagnostico de alzheimer es muy costoso e invasivo y un diagnostico temprano es importante para el cuidado de los pacientes por lo cual se busca explorar el uso de la IA para abaratar los costos del diagnostico de la enfermedad

- **Consideraciones eticas:** Se solicito a los participantes el consentimiento y tambien se cumplio con las medidas de proteccion de datos y eticas exigidas por cada entidad que brinde ayuda o informacion para el desarrollo del proyecto
- **Recoleccion de Datos:** Se utilizaron un total de 70 muestras de videos de 8 a 12 horas interactuando con pacientes en una conversacion amistosa en un ambiente amistoso y familiar, todo esto considerando la forma de actuar y de hablar de las personas con alzheimer, 50 de ellos eran sanos y 20 tenian alzheimer
- **Procesamiento de los datos:** Se recolecto el audio de los videos y se removieron todos los eventos no analizables, dejando como utilizable el 50% y 80% de los datos obtenidos, los primeros de los pacientes con Alzheimer y los segundos los pacientes sanos, los audios se dividieron en segmentos de 60 segundos terminando con un total de 600 segmentos de audio, se utilizó ASSA para obtener informacion auditiva de los audios, se estudio tambien la posibilidad de estudiar los factores emocionales y la dimension fractal Higuchi
- **Modelo:** Se utilizó un Perceptron multicapa con una capa oculta de 100 neuronas y 1000 pasos de entrenamiento junto a Cross validation para lograr optimizar el modelo, el modelo fue alimentado con características extraídas

de los datos utilizando la informacion emocional, los ratios armonicos y en general las características esenciales acusticas

- **Conclusiones:** El modelo fue un exito ya que se logro una precision del 95 %, seguramente debido al debido procesamiento y analisis profundo de las características auditivas a estudiar, el modelo es una muestra de como el conocimiento especializado y los equipos interdisciplinarios pueden lograr y optimizar el adecuado uso de la Inteligencia Artificial en la medicina

4) Artificial Intelligence-Enabled End-To-End Detection and Assessment of Alzheimer's Disease Using Voice

La Inteligencia Artificial presenta muchas ayudas que pueden complementar el diagnostico de enfermedades, el objetivo de este proyecto fue estudiar el uso de las redes pre entrenadas para diagnosticar alzheimer

- **Consideraciones eticas:** Se usaron las consideraciones eticas de la Universidad de Drexel en Filadelfia, USA y se utilizaron datos solicitados de la base de datos de ADReSSo
- **Recoleccion de Datos:** Utilizando los Datos de ADReSSo se recolectaron un total de 247 muestras de audio de las cuales un poco mas de la mitad tenian alzheimer y el resto eran personas sanas
- **Procesamiento de los datos:** Se utilizaron el 70 % de los datos durante el entrenamiento y el 30 % restante se utilizó la validacion, se utilizó Data2Vec para procesar los audios y la informacion extra que contenia la base de datos como el valor de el MMSE para generar vectores con los cuales entrenar el modelo, tambien se utilizó para comparar modelos wav2vec2 utilizando solamente los datos auditivos
- **Modelo:** Se utilizó la libreria de Hugging face para redes neuronales para procesar los datos junto a Cross validation
- **Conclusiones:** Se logro una precision del 73 % utilizando data2vec y un 72 % utilizando wav2vec2, para lograr diagnosticar la enfermedad parece ser que es necesario una mayor cantidad de datos y preprocesamiento de los datos, no obstante el modelo presenta resultados prometedores

5) A deep learning approach to dysphagia-aspiration detecting algorithm through pre- and post-swallowing voice changes

Los costos de diagnosticar Disfagia son muy elevados e incomodos para la mayoría de pacientes por lo cual se busca lograr metodos de diagnostico que sean mas faciles para los pacientes y pueda alcanzar a mas personas por lo cual se busca implementar un modelo que utilice el audio y lo procese rapidamente para obtener los resultados de diagnostico

- **Consideraciones eticas:** Se realizo todo bajo las consideraciones eticas y solicitadas del Hospital Bundang de la Universidad Nacional de Seoul y pidiendo el consentimiento de los pacientes
- **Recoleccion de Datos:** Se recolectaron los datos de 198 pacientes en el hospital entre octubre de 2021 y febrero de 2023 recolectando informacion auditiva antes de consumir alimentos y despues de consumirlos y se utilizaron diferentes dispositivos de recoleccion de audio para evitar sesgos particulares del medio de recoleccion
- **Procesamiento de los datos:** Los audios se limpiaron quitando el ruido, se dividieron en audios de 2 segundos y se procesaron usando las frecuencias de Mel, y

varios procedimientos de audio entre los cuales estaba la transformacion de Fourier de tiempo reducido, se hicieron combinaciones de los fragmentos de antes y despues de consumir alimentos para generar aproximadamente 8000 muestras

- **Modelo:** Se utilizaron redes neuronales preentrenadas con PyTorch, la red fue alimentada con espectrogramas de Mel, se realizaron 3 modelos uno para pacientes femeninos, uno para pacientes masculinos y el otro mixto
- **Conclusiones:** El proposito de esta red fue utilizarla como apoyo al diagnostico de la enfermedad y no como reemplazo, lo cual nos ayuda a reducir los costos de diagnostico en las instituciones, el modelo fue relativamente exitoso con una precision del 81 %

3. METODOLOGIA

Para generar un buen modelo es necesario tener un adecuado procedimiento de recoleccion y analisis de los datos por lo cual es importante comprender y estudiar los alcances de nuestras herramientas. Necesitamos observar que informacion tenemos y cual es posible obtener, en las implementaciones de las ultimas decadas principalmente se han usado bases de datos cerradas con datos diagnosticos y pruebas de audio[6] por lo cual se revisara la disponibilidad de esta informacion y se trabaja con la informacion valiosa recolectada dentro del marco de las posibilidades, ya que la informacion es nuestro mayor aliado para lograr nuestra tarea, ya que sin buenos datos no es posible lograr buenos modelos[10]

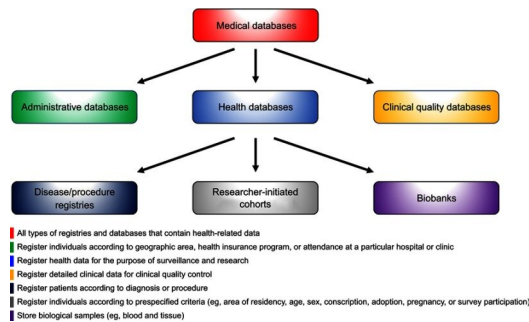


Figura 5. Bases de datos medicas

Para esto se estudiara y buscaran las librerias necesarias, los datos donde entrenar el modelo y se realizaron los respectivos analisis, se realizara el respectivo estudio de los datos

4. ESTRUCTURA DEL PROYECTO

Enfoque y Tecnicas

Este proyecto tiene un enfasis en el estudio del alzheimer, la recoleccion de datos y la estructuracion de un prototipo del modelo, implementando inteligencia artificial para lograr obtener un modelo que con un audio pueda determinar si alguien posee o no alzheimer **Herramientas** Para el proyecto se va a hacer uso de las siguientes Herramientas y Tecnologias:

Github: Donde se Hosteara el acceso al codigo y el Artículo <https://github.com/HanamDavid/AFDA.git>
 Google Colab: Para lograr manipular el codigo y observar la estructura del proyecto Link
 Librerias:

- Tensorflow: Para implementar los algoritmos de machine learning necesario y poder clasificar los datos correctamente y generar procesos utiles
- Sklearn: Para poder dividir los datos del entrenamiento
- Librosa: Pra procesar la informacion de los audios y poder normalizar los datos lo cual nos permitira optimizar y mejorar el rendimiento de la red neuronal y en general con la limpieza de datos
- Matplotlib: Para Visualizar la informacion y asi poder tener una forma de entender los datos con mayor facilidad
- Pandas y Numpy: Para la manipulacion de las Series y los Datos una vez abstraídos de las muestras de Audio

Se utilizo la base de datos **Dementia Bank** la cual incluye una serie de datos medicos de pacientes con demencia y pacientes sanos recolectados especificamente para ML <https://dementia.talkbank.org/ADReSSo-2021/>

1	2	3	4	5	6	7	8	9	10	11	12
name	dementia type	birth	death	first symptom	URLs after symptoms	5 years	5 < 10 years	10 < 15 years			
1	Alzheimer	1910	1985	1975							
2	Alzheimer	1926	2017	2012	https://youtu.be/ThqEIQXQ	https://youtu.be/CdKQDQV8Y8Y	https://youtu.be/7uF6C9QNMNM	https://youtu.be/7uF6C9QNMNM	https://youtu.be/7uF6C9QNMNM	https://youtu.be/7uF6C9QNMNM	https://youtu.be/7uF6C9QNMNM

Figura 6. Datos originales

Esta base de datos contiene un total de 359 muestras de audios recolectados de internet de personas famosas entre las cuales 131 tiene demencia, los audios poseen distintas duraciones y la hay personas que llevan poco tiempo diagnosticadas como pacientes con mas de decada y media con la enfermedad, los datos fueron descargados y subidos a google drive en la carpeta junto al codigo **Procesamiento de Datos**

Primero se descargaron todos los datos y se recortaron a un total de 2 segundos para que fuera mas facil de procesar para la red neuronal, luego se paso por las carpetas para crear un csv que tuviera solamente la informacion redundante ya que el objetivo del proyecto era un clasificador binario, que solamente determinara si el paciente tiene alzheimer o no.

	A	B	C	D	E	F	G	H	I
1	File ID	Class ID	Class Name	Filename	Filepath				
2	0	dementia	TeresaGorman_1	/content/drive/MyDrive/AFDA/Data/dementia/TeresaGorman_1_trimmed.wav					
3	0	dementia	InsMurdoch_10	/content/drive/MyDrive/AFDA/Data/dementia/InsMurdoch_10_trimmed.wav					
4	0	dementia	raphklein_5_1	/content/drive/MyDrive/AFDA/Data/dementia/raphklein_5_1_trimmed.wav					
5	4	0	dementia	RichardOwryn_5	/content/drive/MyDrive/AFDA/Data/dementia/RichardOwryn_5_trimmed.wav				
6	5	0	dementia	RobinWilliams_0	/content/drive/MyDrive/AFDA/Data/dementia/RobinWilliams_0_trimmed.wav				
7	6	0	dementia	peterreven_15	/content/drive/MyDrive/AFDA/Data/dementia/peterreven_15_trimmed.wav				
8	7	0	dementia	DavidProwse_5	/content/drive/MyDrive/AFDA/Data/dementia/DavidProwse_5_trimmed.wav				
9	8	0	dementia	DonaldSterling_10	/content/drive/MyDrive/AFDA/Data/dementia/DonaldSterling_10_trimmed.wav				
10	9	0	dementia	DennisMoore_5	/content/drive/MyDrive/AFDA/Data/dementia/DennisMoore_5_trimmed.wav				
11	10	0	dementia	OmarSharif_0	/content/drive/MyDrive/AFDA/Data/dementia/OmarSharif_0_trimmed.wav				
12	11	0	dementia	jimmclean_10	/content/drive/MyDrive/AFDA/Data/dementia/jimmclean_10_trimmed.wav				
13	12	0	dementia	EstelleGetty_15	/content/drive/MyDrive/AFDA/Data/dementia/EstelleGetty_15_trimmed.wav				
14	13	0	dementia	TonyBennett_0	/content/drive/MyDrive/AFDA/Data/dementia/TonyBennett_0_trimmed.wav				
15	14	0	dementia	DavidProwse_10	/content/drive/MyDrive/AFDA/Data/dementia/DavidProwse_10_trimmed.wav				
16	15	0	dementia	AbeBurrows_5	/content/drive/MyDrive/AFDA/Data/dementia/AbeBurrows_5_trimmed.wav				
17	16	0	dementia	PaulinePhillips_10	/content/drive/MyDrive/AFDA/Data/dementia/PaulinePhillips_10_trimmed.wav				
18	17	0	dementia	StarBowie_5	/content/drive/MyDrive/AFDA/Data/dementia/StarBowie_5_trimmed.wav				
19	18	0	dementia	georgklein_5	/content/drive/MyDrive/AFDA/Data/dementia/georgklein_5_trimmed.wav				
20	19	0	dementia	StellaStevens_1	/content/drive/MyDrive/AFDA/Data/dementia/StellaStevens_1_trimmed.wav				
21	20	0	dementia	EvelynKeyes_10	/content/drive/MyDrive/AFDA/Data/dementia/EvelynKeyes_10_trimmed.wav				
22	21	0	dementia	antonyflew_10	/content/drive/MyDrive/AFDA/Data/dementia/antonyflew_10_trimmed.wav				
23	22	0	dementia	BillBuckner_15	/content/drive/MyDrive/AFDA/Data/dementia/BillBuckner_15_trimmed.wav				
24	23	0	dementia	ErnieSigley_15	/content/drive/MyDrive/AFDA/Data/dementia/ErnieSigley_15_trimmed.wav				
25	24	0	dementia	JimNeidhart_0	/content/drive/MyDrive/AFDA/Data/dementia/JimNeidhart_0_trimmed.wav				
26	25	0	dementia	CaseyKasem_5	/content/drive/MyDrive/AFDA/Data/dementia/CaseyKasem_5_trimmed.wav				

Figura 7. Datos Limpios

Posteriormente ya con las carpetas y el csv, se intercambiaron de posicion cada una de las muestras para facilitar el entrenamiento al aleatorizar el orden de las muestras. Posteriormente se utilizo librosa para extraer las características esenciales de cada uno de los audios con MFCC ya que MFCC permite extraer datos vitales para obtener informacion auditiva de manera mas facil

```
## converting extracted_features to Pandas dataframe
extracted_features_df = pd.DataFrame(extracted_features, columns=['feature', 'class'])
extracted_features_df.head()
```

	feature	class
0	[395.9233, 176.13206, -3.273523, 28.384824, ...]	dementia
1	[221.25842, 140.67139, -20.46207, 37.109334, ...]	dementia
2	[326.48383, 122.79385, -18.557823, 37.252045, ...]	nodementia
3	[271.54138, 72.05179, -13.521994, -4.018288, ...]	nodementia
4	[259.1634, 89.34117, -32.883838, 33.84104, ...]	nodementia

Figura 8. Datos de MFCC

y se guardaron los datos en un vector para el entrenamiento finalmente dividiendo los sets de entrenamiento y de validacion con sklearn, se utilizo el 80% de los datos para el entrenamiento de la red **Modelo**

Para el modelo se uso Tensorflow y se realizaron 3 modelos para poder observar la respuesta de distintos parametros El Primer Modelo fue el siguiente:

```
model = Sequential()
##first layer
model.add(Dense(100, input_shape=(40, 1)))
model.add(Activation('relu'))
model.add(Dropout(0.5))
##second layer
model.add(Dense(200))
model.add(Activation('relu'))
model.add(Dropout(0.5))
##third layer
model.add(Dense(100))
model.add(Activation('relu'))
model.add(Dropout(0.5))
##final layer
model.add(Dense(num_labels))
model.add(Activation('softmax'))
```

~/usr/local/lib/python3.8/dist-packages/keras/src/layers/core/dense.py:87: UserWarning: Do not pass an 'input_shape' / 'input_dim' argument to a layer. When using Sequential no super().__init__(activity_regularizer=activity_regularizer, **kwargs)

Figura 9. Código del Modelo 1

```
model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 100)	4,100
activation (Activation)	(None, 100)	0
dropout (Dropout)	(None, 100)	0
dense_1 (Dense)	(None, 200)	20,200
activation_1 (Activation)	(None, 200)	0
dropout_1 (Dropout)	(None, 200)	0
dense_2 (Dense)	(None, 100)	20,100
activation_2 (Activation)	(None, 100)	0
dropout_2 (Dropout)	(None, 100)	0
dense_3 (Dense)	(None, 2)	202
activation_3 (Activation)	(None, 2)	0

Total params: 44,602 (174.23 KB)
Trainable params: 44,602 (174.23 KB)
Non-trainable params: 0 (0.00 B)

Figura 10. Resumen del Modelo 1

El segundo:

```
from tensorflow.keras.layers import Conv1D, MaxPooling1D, BatchNormalization, Flatten, LSTM

model = Sequential()
model.add(Conv1D(filters=64, kernel_size=3, activation='relu', input_shape=(40, 1)))
model.add(MaxPooling1D(pool_size=2))
model.add(BatchNormalization())
model.add(Conv1D(filters=128, kernel_size=3, activation='relu'))
model.add(MaxPooling1D(pool_size=2))
model.add(BatchNormalization())
model.add(LSTM(100, return_sequences=True))
model.add(Dropout(0.5))
model.add(Flatten())
model.add(Dense(num_labels, activation='softmax'))
```

Figura 11. Código del Modelo 2

```
model.summary()
```

Model: "sequential_5"

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 38, 64)	256
max_pooling1d (MaxPooling1D)	(None, 19, 64)	0
batch_normalization (BatchNormalization)	(None, 19, 64)	256
conv1d_1 (Conv1D)	(None, 17, 128)	24,704
max_pooling1d_1 (MaxPooling1D)	(None, 8, 128)	0
batch_normalization_1 (BatchNormalization)	(None, 8, 128)	512
lstm (LSTM)	(None, 8, 100)	91,600
dropout_15 (Dropout)	(None, 8, 100)	0
flatten (Flatten)	(None, 800)	0
dense_20 (Dense)	(None, 2)	1,602

Total params: 118,930 (464.57 KB)
Trainable params: 118,546 (463.07 KB)
Non-trainable params: 384 (1.50 KB)

Figura 12. Resumen del Modelo 2

y el tercero finalmente:

```
from tensorflow.keras.layers import Conv2D, MaxPooling2D, BatchNormalization, Bidirectional, LSTM
from tensorflow.keras.models import Sequential

# Define the model
model = Sequential()

# Convolutional layers
model.add(Conv2D(32, (3, 3), activation='relu', input_shape=(height, width, channels)))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(BatchNormalization())

model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(BatchNormalization())

model.add(Conv2D(128, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(BatchNormalization())

# Flatten and add LSTM layers
model.add(Flatten())
model.add(Bidirectional(LSTM(64, return_sequences=True)))
model.add(Bidirectional(LSTM(64)))

# Dense layers
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(num_labels, activation='softmax'))
```

Figura 13. Código del Modelo 3

```
model.summary()
```

Model: "sequential_6"

Layer (type)	Output Shape	Param #
conv1d_2 (Conv1D)	(None, 38, 64)	256
max_pooling1d_2 (MaxPooling1D)	(None, 19, 64)	0
batch_normalization_2 (BatchNormalization)	(None, 19, 64)	256
conv1d_3 (Conv1D)	(None, 17, 128)	24,704
max_pooling1d_3 (MaxPooling1D)	(None, 8, 128)	0
batch_normalization_3 (BatchNormalization)	(None, 8, 128)	512
lstm_1 (LSTM)	(None, 8, 100)	91,600
dropout_16 (Dropout)	(None, 8, 100)	0
flatten_1 (Flatten)	(None, 800)	0
dense_21 (Dense)	(None, 2)	1,602

Total params: 118,930 (464.57 KB)
Trainable params: 118,546 (463.07 KB)
Non-trainable params: 384 (1.50 KB)

Figura 14. Resumen del Modelo 3

Se utilizaron diferentes modelos para así lograr determinar como diferentes arquitecturas atacaban el problema, el primer modelo fue un modelo sencillo, mientras que el 2do y 3er modelo fueron un poco mas complejos implementando batch normalization, early stopping, pooling y capas convolucionales y tambien se probaron diferentes valores del batch, al final del

entrenamiento los resultados fueron los siguientes:

- Modelo 2: 89% training y 63% en validacion
- Modelo 3: 77% training y 62% en validacion

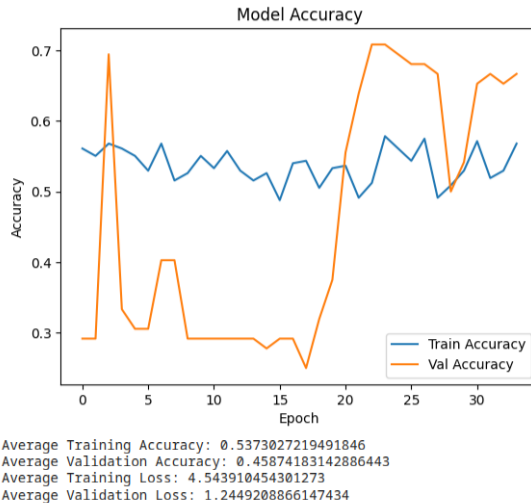


Figura 15. Precision del Modelo 1

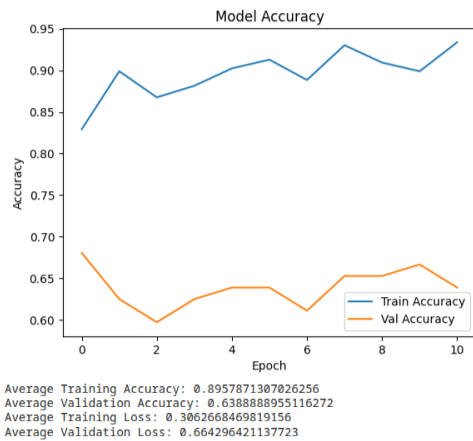


Figura 16. Precision del Modelo 2

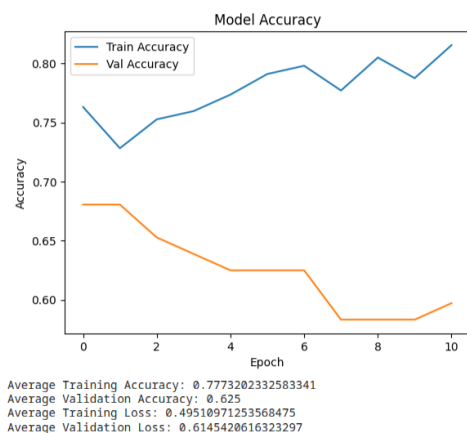


Figura 17. Precision del Modelo 3

Con una precision de los modelos de :

- Modelo 1: 53% training y 45% en validacion

5. RESULTADOS

Se logro estudiar las características particulares del alzheimer y estudiar las consideraciones debidas a la hora de estudiar el diagnostico de enfermedades con Machine Learning, dados los resultados de los 3 modelos, se decidio escoger el Modelo 2 como El ganador debido a que poseia la mayor precision durante el training y durante la validacion. Aun asi los valores siguen de validacion siguen siendo bajos por lo cual es importante explorar a futuro mejores arquitecturas para procesar la informacion, ya que como podemos observar entre el primer modelo y el segundo y tercero hay una diferencia muy notoria de precision y de rendimiento pese a que ambos modelos tomaron un tiempo similar de entrenamiento, la capa convolucional junto a las adiciones al entrenamiento parecen ayudar bastante a mejorar la calidad de prediccion, esto podria a que:

- Las capas convolucionales permiten explorar mas a profundidad características particulares de los datos
- La capa de pooling permite a la red enfocarse en datos mas relevantes y generalizar mejor
- LSTM permite estudiar mejor datos secuenciales de la informacion

Tambien es importante reconocer que en el set de datos escogidos, no todos los pacientes tienen alzheimer si no tambien poseen otros tipos de enfermedades neurologicas, lo cual puede generar que la red no halle todos los patrones particulares del alzheimer es necesario complementar el diagnostico con IA con metodos tradicionales de diagnostico.

Otro factor a considerar es que los datos provienen de la red y no todos poseen la misma calidad y pueden tener mayor o menor cantidad de ruido lo cual puede afectar al modelo, tambien la cantidad de muestras es muy chica, por lo cual aumentando la cantidad de datos para entrenar la red neuronal y obteniendo una mejor bases de datos el modelo podria mejorar bastante. El codigo y el proyecto pueden revisarse en los siguientes links o en los anteriormente listados:

<https://colab.research.google.com/drive/1nOLIXwgonyXJcwL9d8AdiI7VFMG4RBI3unS>

<https://github.com/HanamDavid/AFDA>

6. CONCLUSIONES

El aumento de la capacidad medica para intervenir las enfermedades cambia, se estructura y fomenta buenas practicas medicas dentro de la sociedad. El desarrollo tecnologico es esencial para proteger y cuidar a la poblacion, principalmente a la de alto riesgo, en el caso del alzheimer, cuanto mayor sea la implementacion e investigacion de IA en medicina mayor facilidad abra para mejorar la vida de los pacientes. Esto por supuesto tiene su costo, una gran cantidad de investigacion, un deseo de la sociedad por adoptarlo y una gran cantidad de datos para proporcionarle la informacion adecuada a los modelos. Con el tiempo, la IA toma cada vez mas espacio en nuestra sociedad y al menos en la medicina da buenos prospectos.

Por supuesto es necesario primero priorizar la obtencion moral y etica de datos y de una buena implementacion de sistemas de IoT

para poder recolectar estos datos, ya que sin estos podemos tener modelos con la insuficiente capacidad para detectar la enfermedad o detectarla mal lo cual puede poner en riesgo la vida de muchos pacientes, sin una estructura rígida y moral que sostenga la inteligencia artificial en la medicina no es posible lograr que los modelos sean exitosos y que generen el impacto buscado si no que por el contrario generen falsos positivos, desconfianza de la tecnología y en el peor de los casos la muerte de muchos.

Por último es importante detallar que los avances en Inteligencia artificial brillan sobretodo cuando apoyan a los sistemas tradicionales de investigación y de análisis ya que estos sistemas permiten extrapolar datos y complementar los tratamientos solamente cuando se entiende adecuadamente las circunstancias donde es aplicable, por lo cual es necesario que todo avance este siendo revisado con equipos interdisciplinarios para así poder explotar el uso de la tecnología en la comunidad científica y en este caso en la comunidad médica

7. REFERENCIAS BIBLIOGRAFICAS

REFERENCIAS

- [1] "Alzheimer's disease: early diagnosis and treatment," HKMJ, Jan. 17, 2015. <https://www.hkmj.org/abstracts/v18n3/228.htm>
- [2] R. L. Ernst and J. W. Hay, "The US economic and social costs of Alzheimer's disease revisited.," American Journal of Public Health, vol. 84, no. 8, pp. 1261–1264, Aug. 1994, doi: 10.2105/ajph.84.8.1261.
- [3] K. Qian, B. Hu, Y. Yamamoto, and B. W. Schuller, "The Voice of the Body: Why AI Should Listen to It and an archive.," Cyborg and Bionic Systems, vol. 4, Jan. 2023, doi: 10.34133/cbsystems.0005.
- [4] F. Agbavor and H. Liang, "Artificial Intelligence-Enabled End-To-End detection and assessment of Alzheimer's disease using Voice.," Brain Sciences, vol. 13, no. 1, p. 28, Dec. 2022, doi: 10.3390/brainsci13010028.
- [5] H. Lin et al., "Identification of digital voice biomarkers for cognitive health.," Exploration of Medicine, vol. 1, no. 6, pp. 406–417, Dec. 2020, doi: 10.37349/emed.2020.00028.
- [6] E. Evangelista et al., "Current Practices in voice data collection and Limitations to Voice AI Research: A national survey.," The Laryngoscope, Dec. 2023, doi: 10.1002/lary.31052.
- [7] M. Chen and M. Décary, "Artificial intelligence in healthcare: An essential guide for health leaders.," Healthcare Management Forum, vol. 33, no. 1, pp. 10–18, Sep. 2019, doi: 10.1177/0840470419873123.
- [8] A. Baevski, W.-N. Hsu, Q. Xu, A. Babu, J. Gu, and M. Auli, "data2vec: A General Framework for Self-supervised Learning in Speech, Vision and Language.," PMLR, Jun. 28, 2022. <https://proceedings.mlr.press/v162/baevski22a.html>
- [9] D. Berrar, "Cross-Validation.," in Elsevier eBooks, 2019, pp. 542–545. doi: 10.1016/b978-0-12-809633-8.20349-x.
- [10] Womersley K, Fulford KB, Peile E, Korulus P, Handa A. Hearing the patient's voice in AI-enhanced healthcare. BMJ. 2023 Nov 21;383:2758. doi: 10.1136/bmj.p2758. PMID: 37989502.
- [11] M. Alzubaidi et al., "The role of neural network for the detection of Parkinson's Disease: A scoping review.," Healthcare, vol. 9, no. 6, p. 740, Jun. 2021, doi: 10.3390/healthcare9060740.
- [12] J. Laguarda, F. Hueto, P. Rajasekaran, S. E. Sarma, and B. Subirana, "Longitudinal Speech biomarkers for automated Alzheimer's detection.," Research Square (Research Square), Aug. 2020, doi: 10.21203/rs.3.rs-56078/v1.
- [13] J. Laguarda, F. Hueto, and B. Subirana, "COVID-19 Artificial intelligence diagnosis using only cough recordings.," IEEE Open Journal of Engineering in Medicine and Biology, vol. 1, pp. 275–281, Jan. 2020, doi: 10.1109/ojemb.2020.3026928.
- [14] K. L. De Ipiña et al., "On the Selection of Non-Invasive Methods Based on Speech Analysis Oriented to Automatic Alzheimer Disease Diagnosis.," Sensors (Basel), vol. 13, no. 5, pp. 6730–6745, May 2013, doi: 10.3390/s130506730.
- [15] J. Huang, P. Guo, S. Zhang, M. Ji, and R. An, "Use of Deep Neural Networks to Predict Obesity with Short Audio Recordings: A Pilot Study (Preprint).," JMIR AI, Nov. 2023, doi: 10.2196/54885.
- [16] J.-M. Kim, M.-S. Kim, S.-Y. Choi, K. Lee, and J. S. Ryu, "A deep learning approach to dysphagia-aspiration detecting algorithm through pre- and post-swallowing voice changes.," Frontiers in Bioengineering and Biotechnology, vol. 12, Aug. 2024, doi: 10.3389/fbioe.2024.1433087.
- [17] UCSF, "AI and Alzheimer's Research.," YouTube. May 10, 2022. [Online]. Available: <https://www.youtube.com/watch?v=zkvKtER5tI>