



UNIVERSIDAD DE GUAYAQUIL

FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS

CARRERA DE INGENIERÍA EN SISTEMAS

COMPUTACIONALES

ALGORITMO PARA DIAGNÓSTICO DE

TRASTORNO DISLÁLICOS EN

PACIENTES A PARTIR DE LA

FICHA LOGOPÉDICA

PROYECTO DE TITULACIÓN

Previa a la obtención del Título de:

INGENIERO EN SISTEMAS COMPUTACIONALES

AUTORES:

PHILCO BAQUE JORGE FERNANDO

SERRANO GUAMÁN JOSÉ MIGUEL

TUTOR:

ING. REYES WAGNIO MANUEL

GUAYAQUIL – ECUADOR

2021

  		
REPOSITORIO NACIONAL EN CIENCIAS Y TECNOLOGÍAS		
FICHA DE REGISTRO DE TRABAJO DE TITULACIÓN		
TÍTULO: <i>Algoritmo para diagnóstico de trastorno dislálico en pacientes a partir de la ficha logopédica</i>		
AUTOR(ES): Jorge Fernando Philco Baque José Miguel Serrano Guamán		REVISOR: Ing. Segundo Delgado Menoscal, PhD.
INSTITUCIÓN: Universidad de Guayaquil		FACULTAD: Ciencias Matemáticas y Físicas
CARRERA: Ingeniería en Sistemas Computacionales		
FECHA DE PUBLICACIÓN:		N° DE PÁGS.: 217
ÁREA TEMÁTICA: INVESTIGACIÓN		
PALABRAS CLAVES: FICHAS LOGOPEDICAS, ALGORITMO DE DIAGNOSTICO BASADO EN REGLAS, TRASTORNO DISLALICO, DISLALIA		
<p>RESUMEN: La Dislalia funcional se define como una alteración del lenguaje, la cual se suele presentar con gran frecuencia en infantes de 5 años de edad en adelante y en la variedad de los casos este problema suele permanecer hasta la edad adulta si no es tratada en su debido momento, las dificultades que se acarrearán con este tipo de trastornos de lenguaje desencadenan severas consecuencias de las cuales podemos mencionar el bullying, discriminación, dificultad de aprendizaje, para obtener un diagnóstico de los pacientes que padecen estos trastornos se procedió al desarrollo de un algoritmo para diagnóstico de dislalia funcional, la cual tendrá como objetivo primordial leer las entradas del dataset el cual contendrá las fichas logopédicas realizadas a los pacientes y así poder determinar en qué fonema y en qué posición radica estos problemas de trastornos. Esta necesidad de implementación radica porque en Guayaquil existe muy pocos centros terapéuticos que ayuden a tratar este tipo de trastornos y por su alto costo de consulta lo cual impiden que los niños culminen los tratamientos empezados. Para el desarrollo del algoritmo se realizaron entrevistas a especialistas logopedas y así mismo se recopiló datos de las fichas logopédicas para poderlas adaptar y así tener un diagnóstico correcto, también se utilizó el lenguaje de programación Python 3.9 y según las investigaciones realizadas y los algoritmos seleccionados, el mejor posicionado y de más sencilla aplicación resultó el Basado en Reglas, con él se obtuvo una gran precisión en comparación de los demás algoritmos consultados. Finalmente se muestran las conclusiones donde indican los resultados obtenidos por parte del algoritmo para diagnóstico de dislalia funcional, en donde a partir de estos resultados el especialista logopeda puede tener una noción más clara y así poder decidir cuál sea el tratamiento que se le pueda aplicar al paciente evaluado.</p>		
N° DE REGISTRO:		N° DE CLASIFICACIÓN:
DIRECCIÓN URL:		
ADJUNTO PDF	SI <input checked="" type="checkbox"/>	NO <input type="checkbox"/>
CONTACTO CON AUTOR(ES):	Teléfono: 0993664319 0989011387	Email: jorge.philcob@ug.edu.ec jose.serranog@ug.edu.ec
CONTACTO DE LA INSTITUCIÓN	Nombre: Ab. Juan Chávez Atocha	
	Teléfono: 2307729	
	Email: juan.chaveza@ug.edu.ec	

APROBACIÓN DEL TUTOR

En mi calidad de Tutor(a) del Trabajo de Titulación, “ALGORITMO PARA DIAGNOSTICO DE TRASTORNOS DISLALICOS EN PACIENTES A PARTIR DE LA FICHA LOGOPEDICA” elaborado por los Sres. JORGE FERNANDO PHILCO BAQUE y JOSÉ MIGUEL SERRANO GUAMÁN, **estudiantes no titulados** de la Carrera de Ingeniería en Sistemas Computacionales, Facultad de Ciencias Matemáticas y Físicas de la Universidad de Guayaquil, previo a la obtención del Título de Ingeniero(a) en Sistemas Computacionales, me permito declarar que luego de haber orientado, estudiado y revisado, la **apruebo** en todas sus partes.

Atentamente,

Ing. Manuel Reyes Wagnio

TUTOR

DEDICATORIA

Es mi deseo dedicar el presente proyecto a mi madre, mis hermanas, a mi querida esposa, a mi pequeño hijo y en especial a la Lcda. Natividad Guerrero por su apoyo incondicional durante todos estos años de estudio, ellos son la clave para lograr culminar con éxitos mi carrera universitaria.

Jorge Fernando Philco Baque

Es mi deseo el dedicar este proyecto de titulación a mis padres y mi hermana, quienes han sido mi mayor motivación para que pueda cumplir con todas mis metas, a mi familia, a mi novia, a mis amigos y demás personas que siempre estuvieron presentes en el transcurso de este logro.

José Serrano Guamán

AGRADECIMIENTO

En primer lugar, agradezco a Dios, en segundo lugar, a mi familia por su apoyo incondicional siempre, a mis amigos por brindarme la motivación necesaria para seguir adelante, en tercer lugar, al Ing. Manuel Reyes Wagnio, por su guía durante todo este proceso.

Jorge Fernando Philco Baque

En primer lugar, agradezco a Dios por guiarme en todo momento, en segundo a mis padres, por su buena enseñanza y confiar siempre en mí y finalmente muchas gracias al Ing. Manuel Reyes, quien fue la mayor guía para que todo esto sea posible y por todo el apoyo brindado. Gracias a todos por eso y por muchos más.

José Miguel Serrano Guamán

TRIBUNAL PROYECTO DE TITULACIÓN

Ing. Douglas Iturburu Salvador, M.Sc.
DECANO DE LA FACULTAD
CIENCIAS MATEMÁTICAS Y FÍSICAS

Ing. Lorenzo Cevallos Torres, Mgs.
DIRECTOR DE LA CARRERA DE
INGENIERÍA EN SISTEMAS
COMPUTACIONALES

Ing. Manuel Reyes Wagnio
PROFESOR TUTOR DEL PROYECTO
DE TITULACIÓN

Ing. Segundo Delgado Menoscal, PhD.
PROFESOR REVISOR DEL PROYECTO
DE TITULACIÓN

Ab. Juan Chávez Atocha, Esp.
SECRETARIO

DECLARACIÓN EXPRESA

“La responsabilidad del contenido de este Proyecto de Titulación, me corresponden exclusivamente; y el patrimonio intelectual de la misma a la UNIVERSIDAD DE GUAYAQUIL”.

JORGE FERNANDO PHILCO BAQUE

JOSÉ MIGUEL SERRANO GUAMÁN



CESIÓN DE DERECHOS DE AUTOR

Ingeniero

Douglas Iturburu Salvador, M.Sc.

DECANO DE LA FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS

Presente.

A través de este medio indico a usted que procedo a realizar la entrega de la cesión de derechos de autor en forma libre y voluntaria del trabajo de titulación “**Algoritmo para diagnóstico de trastorno dislábico en pacientes a partir de la ficha logopédica**”, realizado como requisito previo para la obtención del Título de Ingeniero(a) en Sistemas Computacionales de la Universidad de Guayaquil.

Guayaquil, Septiembre de 2021.

Jorge Fernando Philco Baque
C.I. N° 0922975149

José Miguel Serrano Guamán
C.I. N° 0705966935

CERTIFICADO DE ACEPTACIÓN DEL TUTOR

En mi calidad de Tutor del Proyecto de Titulación, nombrado por el Consejo Directivo de la Facultad de Ciencias Matemáticas y Físicas de la Universidad de Guayaquil.

CERTIFICO:

Que he analizado el Proyecto de Titulación presentado por el/la/los estudiante(s) **JORGE FERNANDO PHILCO BAQUE, JOSÉ MIGUEL SERRANO GUAMÁN**, como requisito previo para optar por el Título de Ingeniero en Sistemas Computacionales cuyo proyecto es:

ALGORITMO PARA DIAGNOSTICO DE TRASTORNOS DISLALICOS EN PACIENTES A PARTIR DE LA FICHA LOGOPEDICA.

Considero aprobado el trabajo en su totalidad.

Presentado por:

Philco Baque Jorge Fernando

Cédula de identidad N°
0922975149

Serrano Guamán José Miguel

Cédula de identidad N°
0705966935

Tutor: _____

Ing. Manuel Reyes Wagnio

Guayaquil, septiembre de 2021



**UNIVERSIDAD DE GUAYAQUIL
FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES**

AUTORIZACIÓN PARA PUBLICACIÓN DE PROYECTO DE TITULACIÓN EN FORMATO DIGITAL

1. Identificación del Proyecto de Titulación

Nombre del Estudiante: Jorge Fernando Philco Baque	
Dirección: Coop. Juan Montalvo Lotes Alegría	
Teléfono 0993664319	Email: jorge.philcob@ug.edu.ec

Nombre del Estudiante: José Miguel Serrano Guamán	
Dirección: Av. San Martín e/ Junín y Tarquí	
Teléfono: 0989011387	Email: jose.serranog@ug.edu.ec

Facultad: Ciencias Matemáticas y Físicas
Carrera: Ingeniería en Sistemas Computacionales
Proyecto de Titulación al que opta: Investigación
Profesor Tutor: Ing. Manuel Reyes Wagnio

Título del Proyecto de Titulación: Algoritmo para diagnóstico de trastornos dislálicos en pacientes a partir de la ficha logopédica.

Palabras Claves: Fichas logopédicas, algoritmos de diagnóstico, basado en reglas, trastorno dislálico, dislalia
--

2. Autorización de Publicación de Versión Electrónica del Proyecto de Titulación

A través de este medio autorizo a la Biblioteca de la Universidad de Guayaquil y a la Facultad de Ciencias Matemáticas y Físicas a publicar la versión electrónica de este Proyecto de Titulación.

Publicación Electrónica:

Inmediata	X	Después de 1 año
-----------	----------	------------------

Firma Estudiante:

Philco Baque Jorge Fernando

0922975149
Cédula de identidad N°

Serrano Guamán José Miguel

0705966935
Cédula de identidad N°

3. Forma de envío:

El texto del Proyecto de Titulación debe ser enviado en formato Word, como archivo .docx, .RTF o Puf para PC. Las imágenes que la acompañen pueden ser: .gif, .jpg o .TIFF.

DVDROM

CDROM

ÍNDICE GENERAL

FICHA DE REGISTRO DE TRABAJO DE TITULACIÓN	ii
APROBACIÓN DEL TUTOR.....	iii
DEDICATORIA.....	iv
AGRADECIMIENTO	v
DECLARACIÓN EXPRESA.....	vii
CESIÓN DE DERECHOS DE AUTOR	viii
CERTIFICADO DE ACEPTACIÓN DEL TUTOR	ix
AUTORIZACIÓN PARA PUBLICACIÓN DE PROYECTO DE TITULACIÓN EN FORMATO DIGITAL	x
ÍNDICE GENERAL	xi
ÍNDICE DE TABLAS.....	xix
ÍNDICE DE FIGURAS.....	xxi
ABREVIATURAS.....	xxiii
RESUMEN.....	xxiv
ABSTRACT.....	xxv
INTRODUCCIÓN	1
CAPÍTULO I.....	4
PLANTEAMIENTO DEL PROBLEMA	4

Descripción de la situación problemática	5
Ubicación del problema en un contexto.....	5
Situación conflicto nudos críticos.....	6
Delimitación del problema.....	7
Evaluación del Problema	7
Causas y consecuencias del problema	9
Formulación del problema	10
Objetivos del proyecto	10
Objetivo general.....	10
Objetivos específicos	10
Alcance del proyecto	11
Justificación e importancia	12
Limitaciones del estudio	13
CAPÍTULO II	14
MARCO TEÓRICO	14
Antecedentes del estudio.....	14
Investigaciones a nivel Internacional.....	14
Investigación a nivel Nacional.....	15
Investigación a nivel local	16
Fundamentación teórica.....	17

Dislalia	17
Dislalia Fisiológica o evolutiva	17
Dislalia Audiógena.....	18
Dislalia Orgánica	18
Dislalia Funcional	18
Dislalia Funcional	19
Déficit fonético y fonológico	19
Clasificación según el déficit.....	19
Factores etiológicos	20
Detección de la dislalia funcional	21
Tratamiento dislalia funcional	22
Ficha logopédica	23
Historia Clínica	23
Ficha de articulación – pronunciación	23
Ficha exploración de los órganos articulatorios y de fonación.....	25
Exploración del lenguaje expresivo	26
Exploración del lenguaje impresivo.....	27
Inteligencia artificial	27
Algoritmos	28
Definición	28

Clasificación de los algoritmos	29
Algoritmos Bayesianos	29
Algoritmo arboles de decisión	35
Algoritmos de regresión.....	39
Algoritmos de agrupamiento.....	43
Algoritmos basados en Instancia	45
Basado en reglas	46
Análisis de Algoritmos.....	47
1. Definición de objetivos e indicadores claves.....	47
2. Identificar los elementos a comparar	48
3. Levantamiento de datos	49
4. Tabla comparativa.....	59
5. Interpretación de resultados y conclusión.....	61
Herramientas de desarrollo de software.....	62
Python 3.9	62
Comparación de Python frente a otros lenguajes.....	63
STAT::FIT	64
@Risk 8.2	65
SPSS.....	65
Revisiones sistemáticas	65

Metaanálisis	66
Diseño del metaanálisis.....	66
Instrumentos utilizados para la recolección de datos.....	66
Pregunta científica a contestarse	67
Definiciones conceptuales.....	67
Aprendizaje supervisado.....	67
Diagnóstico	68
Fonema.....	68
Inteligencia Artificial	68
Logopeda.....	68
Machine Learning	69
Python	69
Trastornos	69
CAPÍTULO III.....	70
METODOLOGÍA DE LA INVESTIGACIÓN	70
Tipo de investigación	71
Diseño metodológico de la investigación.....	71
Metodología de investigación.....	72
Población y muestra.....	74
Población.....	74

Población Objetivo.....	76
Muestra	76
Muestreo	77
Marco Muestral.....	78
Procesamiento y análisis.....	78
Técnicas de recolección de datos.....	78
Técnicas estadísticas para el procesamiento de la información.....	80
Metodología de desarrollo del prototipo.....	92
Metodología Kanban.....	92
<i>Fases de la metodología</i>	93
Fase 1	93
Fase 2	94
Fase 3	98
Pruebas Inducidas	109
Beneficiarios directos e indirectos del proyecto	115
Beneficiarios Directos	115
Beneficiarios Indirectos	115
Entregables del proyecto	115
Propuesta	116
Criterios de validación de la propuesta	117

Informe del plan de pruebas	117
Juicio de experto	119
Análisis de datos	120
Resultados	125
CAPÍTULO IV	127
CONCLUSIONES Y RECOMENDACIONES	127
Conclusiones	127
Recomendaciones	128
Trabajos futuros	128
REFERENCIAS BIBLIOGRÁFICAS	129
ANEXOS	145
Anexo 1. Planificación de actividades del proyecto	145
Anexo 2. Geo-localización del problema.....	146
Anexo 3. Fundación Legal.....	147
Anexo 4. Criterios éticos a utilizarse en el desarrollo del proyecto.....	150
Anexo 5. Formatos de técnicas de recolección de datos.....	152
Anexo 6. Validación de expertos.	157
Anexo 7. Tabla del Meta-análisis	166
Anexo 8. Acta de entrega y recepción	170
Anexo 9. Artículo científico	173

1. Introducción	173
1.1. Trabajos Relacionados.....	174
2. Metodología	174
3. Materiales y Métodos.....	175
4. Caso de estudio.....	177
5. Resultados.....	184
6. Conclusiones.....	184
7. Referencias.....	184

ÍNDICE DE TABLAS

Tabla 1 Delimitación del problema.....	7
Tabla 2 Matriz de causas y consecuencias del problema.....	9
Tabla 3 Fenómeno Lingüístico y fonemas alterados.	20
Tabla 4 Ficha técnica ELA-r.....	24
Tabla 5 Ficha técnica EDAF.....	26
Tabla 6 Estudio comparativo herramientas de desarrollo de algoritmos.....	49
Tabla 7 Características de Hardware y Software.....	51
Tabla 8 Software Disponible.....	52
Tabla 9 Hardware y software de algoritmo basado en reglas.....	53
Tabla 10 Características de Hardware y Software.....	54
Tabla 11 Tiempos de ejecución K-NN.....	58
Tabla 12 Resultados en función de los indicadores vs los resultados de los algoritmos.....	59
Tabla 13 Comparación entre Python y R.....	63
Tabla 14 Fuentes bibliográficas.....	66
Tabla 15 Hospitales o Instituciones de especialistas del lenguaje.....	75
Tabla 16 Población de la ciudad de Guayaquil para entrevista.....	77
Tabla 17 Pregunta 1.....	80
Tabla 18 Pregunta 2.....	82
Tabla 19 Pregunta 3.....	83
Tabla 20 Pregunta 4.....	84
Tabla 21 Pregunta 5.....	86

Tabla 22 Pregunta 6	87
Tabla 23 Pregunta 7	88
Tabla 24 Pregunta 8	89
Tabla 25 Pregunta 9	91
Tabla 26 Librerías dentro del algoritmo	99
Tabla 27 Variables dentro del algoritmo.....	99
Tabla 28 Variables para información visualización.....	100
Tabla 29 Variables dentro de la función diagnosticar.....	100
Tabla 30 Tipos de variables	101
Tabla 31 Funciones dentro del algoritmo para diagnostico	102
Tabla 32 Métodos dentro del algoritmo.....	103
Tabla 33 Condicionales.....	104
Tabla 34 Bucles dentro del código.....	105
Tabla 35 Procesos dentro del algoritmo.....	105
Tabla 36 Entradas y salidas del algoritmo	106
Tabla 37 Operacionalización de variables	108
Tabla 38 Factor tiempo y espacio algoritmo basado en reglas en máquina de investigadores..	114
Tabla 39 Clasificación de déficit de fonemas.	116
Tabla 40 Información de Expertos.....	120
Tabla 41 Resumen del procesamiento de los casos: Pregunta 4 y 8.....	121
Tabla 42 Tabla de contingencia pregunta 4 y 8	122
Tabla 43 Pruebas de chi-cuadrado	123
Tabla 44 Resumen del procesamiento de los casos	124

ÍNDICE DE FIGURAS

Figura 1 Precisión Naive bayes	55
Figura 2 Precisión algoritmo K-NN.....	56
Figura 3 Precisión algoritmo árboles de decisión.....	56
Figura 4 Tiempo Ejecución Naive Bayes	57
Figura 5 Tiempo Ejecución basado en reglas	59
Figura 6 Línea de tendencia comparativa entre, Python, R y C	64
Figura 7 Pregunta 1: Análisis gráfico de la pregunta número 1 de la entrevista	81
Figura 8 Pregunta 2: Análisis gráfico de la pregunta número 2 de la entrevista.	82
Figura 9 Pregunta 3: Análisis gráfico de la pregunta número 3 de la entrevista.	83
Figura 10 Pregunta 4: Análisis gráfico de la pregunta número 4 de la entrevista.	85
Figura 11 Pregunta 5: Análisis gráfico de la pregunta número 5 de la entrevista.	86
Figura 12 Pregunta 6: Análisis gráfico de la pregunta número 6 de la entrevista.	87
Figura 13 Pregunta 7: Análisis gráfico de la pregunta número 7 de la entrevista	89
Figura 14 Pregunta 8: Análisis gráfico de la pregunta número 8 de la entrevista.	90
Figura 15 Pregunta 9: Análisis gráfico de la pregunta número 9 de la entrevista	91
Figura 16 STAT::FIT. Distribución de la variable M.....	95
Figura 17 Simulación de datos. Microsoft Visual Basic.	96
Figura 18 Iteraciones por cada una de la clasificación de déficit de fonemas.....	97
Figura 19 Simulación Montecarlo. Resultado final de la simulación.....	97
Figura 20 Proceso de implementación del algoritmo	106
Figura 21 Implementación de algoritmo.....	107

Figura 22 Pantalla de visualización del algoritmo en ejecución.....	107
Figura 23 Pantalla de visualización del diagnóstico al paciente uno.....	108
Figura 24 Prueba N° 1 Dataset con datos de pacientes a diagnosticar	109
Figura 25 Salida por consola del diagnóstico paciente 1	110
Figura 26 Salida por consola del diagnóstico paciente 3.....	110
Figura 27 Prueba N° 2 Dataset con datos de pacientes a diagnosticar	111
Figura 28 Salida por consola del diagnóstico paciente 2.....	111
Figura 29 Salida por consola del diagnóstico paciente 4.....	112
Figura 30 Prueba N° 3 Dataset con datos de pacientes a diagnosticar	112
Figura 31 Salida por consola del diagnóstico del paciente 5.....	113

ABREVIATURAS

AD	Árbol de decisión
BN	Bayesian Network
FCMF	Facultad de Ciencias Matemáticas y Físicas
ELA	Examen Logopédico de Articulación
ELA-R	Examen Logopédico de Articulación - revisado
ELCE	Exploración del lenguaje comprensivo y expresivo
GNB	Gaussian Naive Bayes
IA	Inteligencia Artificial
Ing.	Ingeniero
K-NN	K-Nearest Neighbor
MNB	Multinomial Naive Bayes
M.Sc.	Máster
PLN	Procesamiento del Lenguaje Natural
SNC	Sistema Nervioso Central
UG	Universidad de Guayaquil
URL	Localizador de Fuente Uniforme
WWW	World Wide Web (Red Mundial)



UNIVERSIDAD DE GUAYAQUIL
FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES
 ALGORITMO PARA DIAGNÓSTICO DE
 TRASTORNOS DISLÁLICOS EN
 PACIENTES A PARTIR DE
 LA FICHA LOGOPÉDICA

Autor(es): Jorge Fernando Philco Baque
 C.I. N° 0922975149
 José Miguel Serrano Guamán
 C.I. N° 0705966935

Tutor: Ing. Manuel Reyes Wagnio

RESUMEN

La Dislalia funcional se define como una alteración del lenguaje, la cual se suele presentar con gran frecuencia en infantes de 5 años de edad en adelante y en la variedad de los casos este problema suele permanecer hasta la edad adulta si no es tratada en su debido momento, las dificultades que se acarrearán con este tipo de trastornos de lenguaje desencadenan severas consecuencias de las cuales podemos mencionar el bullying, discriminación, dificultad de aprendizaje, para obtener un diagnóstico de los pacientes que padecen estos trastornos se procedió al desarrollo de un algoritmo para diagnóstico de dislalia funcional, la cual tendrá como objetivo primordial leer las entradas del dataset el cual contendrá las fichas logopédicas realizadas a los pacientes y así poder determinar en qué fonema y en qué posición radica estos problemas de trastornos. Esta necesidad de implementación radica porque en Guayaquil existe muy pocos centros terapéuticos que ayuden a tratar este tipo de trastornos y por su alto costo de consulta lo cual impiden que los niños culminen los tratamientos empezados. Para el desarrollo del algoritmo se realizaron entrevistas a especialistas logopedas y así mismo se recopiló datos de las fichas logopédicas para poderlas adaptar y así tener un diagnóstico correcto, también se utilizó el lenguaje de programación Python 3.9 y según las investigaciones realizadas y los algoritmos seleccionados, el mejor posicionado y de más sencilla aplicación resultó el Basado en Reglas, con él se obtuvo una gran precisión en comparación de los demás algoritmos consultados. Finalmente se muestran las conclusiones donde indican los resultados obtenidos por parte del algoritmo para diagnóstico de dislalia funcional, en donde a partir de estos resultados el especialista logopeda puede tener una noción más clara y así poder decidir cuál sea el tratamiento que se le pueda aplicar al paciente evaluado.

Palabras clave: FICHAS LOGOPÉDICAS, ALGORITMO DE DIAGNÓSTICO, BASADO EN REGLAS, TRASTORNO DISLALICO, DISLALIA.



UNIVERSIDAD DE GUAYAQUIL
FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES
ALGORITMO PARA DIAGNÓSTICO DE
TRASTORNOS DISLÁLICOS EN
PACIENTES A PARTIR DE
LA FICHA LOGOPÉDICA

Author(s): Jorge Fernando Philco Baque
C.I. N° 0922975249
José Miguel Serrano Guamán
C.I. N° 0705966935

Tutor: Ing. Manuel Reyes Wagnio

ABSTRACT

Dyslalia is defined as a language alteration, which usually occurs with great frequency in infants with an age range between 5 to 12 years and in the variety of cases this problem usually remains until adulthood if it is not treated in due course, the difficulties that are brought about with this type of language disorders trigger severe consequences of which we can mention bullying, discrimination, learning difficulty, in order to obtain a diagnosis of patients suffering from these disorders, we proceeded to the development of a dyslalia diagnosis algorithm, which will have as its main objective to read the dataset entries which will contain the speech therapy files made to the patients and thus be able to determine in which phoneme and in which position these dyslalia problems lie. This need for implementation is due to the fact that in Guayaquil there are very few therapeutic centers that help treat these types of disorders and because of their high cost of consultation, which prevents children from completing the treatments they have started. For the development of the algorithm, interviews were conducted with speech therapist specialists and data from the speech therapy files were also collected to be able to adapt them and thus have a correct diagnosis, the Python programming language was also used and according to the research carried out and the selected algorithms, the rule-based one was better positioned and easier to apply, with it a great precision was obtained compared to the other algorithms consulted. Finally, the conclusions are shown where they indicate the results obtained by the dyslalia diagnosis algorithm, where from these results the speech therapist specialist can have a clearer notion and thus be able to decide what treatment can be applied to the patient evaluated.

Key words: SPEECH THERAPY FILES, DIAGNOSIS ALGORITHM, RULE-BASED, DYSLALIC DISORDER, DYSLALIA.

INTRODUCCIÓN

En este trabajo investigativo se busca informar de las alteraciones fonéticas (Dislalia) que se presentan con mayor frecuencia en infantes de temprana edad lo cual acarrea mayores consecuencias ya sea en el ámbito educativo como en su desarrollo personal, en la mayoría de los casos estos problemas no son corregidos de una manera correcta lo que ha conllevado a que las personas que lo padecen sean víctimas de bullying, ocasionándoles complicaciones hasta su edad adulta.

La Dislalia es una alteración fonética la cual consiste en la dificultad de articular, cambiar, omitir o sustituir diferentes fonemas de forma inapropiada, cabe recalcar que existen diversos factores que intervienen en los diferentes tipos de dislalia entre los cuales podemos mencionar: orgánicos, evolutivos y funcionales, siendo este último factor principal por el cual se desarrolló el presente proyecto.

Es importante aclarar que uno de los tipos de dislalia más frecuentes que afectan a los infantes es la denominada dislalia funcional, esta se puede reconocer por el uso inadecuado de los órganos buco fonatorios para la pronunciación de los sonidos, se puede decir que estas anomalías se las puede apreciar en los infantes con edad superior a 5 años, los principales factores que la ocasionan son el entorno donde se desenvuelve el infante, también una causa principal es que los infantes reciben demasiado consentimiento por parte de sus familiares y no lo ven como un problema real y no es tratado a tiempo con un especialista.

Por otra parte, toda persona que padezca de este tipo de trastorno tiene que dirigirse a un centro de salud para ser tratados por profesionales especializados en estos problemas de trastorno dislállico, los cuales se los denomina logopedas, estos realizan sus tratamientos basados en

terapias con diversos ejercicios, utilizando los test de evaluación ELA-R, PAF las cuales se denominan fichas logopédicas, donde el infante repite varias palabras que contienen el fonema a evaluar, para poder diagnosticar en cuál de estas tienen mayor grado de dificultad en su pronunciación y así poder corregir y erradicar este problema en sus pacientes.

Gracias a las herramientas y avances tecnológicos con los que se cuenta actualmente, brindan la apertura para poder realizar este proyecto el cual tiene como finalidad desarrollar un prototipo de algoritmo para diagnóstico de dislalia funcional a partir de las fichas logopédicas, donde una vez obtenido los resultados, el algoritmo lo que realizará es brindar un diagnóstico para conocer en cuales fonemas presenta dificultad el paciente. A continuación, se les mencionará la estructura de los diferentes capítulos:

En el **Capítulo I**, se indica un punto importante lo cual es la problemática, en donde se analizan los nudos críticos, también se da a conocer las dificultades que presentan los pacientes con dislalia funcional en Ecuador, al no contar con gran cantidad de especialistas que ayuden a diagnosticar estos trastornos, también se establecen los objetivos tanto general como específicos, se presentan los alcances del proyecto y se detalla la justificación e importancia de este.

En el **Capítulo II**, contiene toda la fundamentación teórica del proyecto, la cual tiene información relevante, ideas necesarias, lo cual aportara a la solución del problema suscitado, se definen los tipos de dislalia existentes, los antecedentes de estudio que aportaran y serán la guía que se tomen en cuenta para la realización del algoritmo para diagnóstico.

En el **Capítulo III**, Aquí se detalla la población, muestra, también se define la metodología usada para el desarrollo del prototipo de algoritmo, como las diferentes técnicas para recopilar la información, se muestra el análisis obtenido sobre la entrevista efectuada a los especialistas logopedas que evalúen el grado de aceptación para el desarrollo del algoritmo.

Además, se hace una validación de juicio de expertos que evalúen a través de criterios, la eficiencia de los resultados mostrados en el algoritmo.

En el **Capítulo IV**, finalmente en este apartado se presenta todos los resultados obtenidos, incluyendo las conclusiones y recomendaciones del trabajo realizado, bibliografías y anexos tales como: Capturas de pantallas de entrevistas, porciones de código diseñado, artículo científico y así poder ayudar a trabajos futuros que se relacionen con el mismo desarrollo de trastornos dislálicos.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

El diagnóstico temprano y el principio atinado de la administración de los casos del trastorno de la articulación de los fonemas (Dislalia) son importantes para el decrecimiento de esta patología en la población infantil. La falta de algoritmos de diagnóstico con correcta precisión basados en un análisis cuantitativo de la transformación de un análisis cualitativo de un especialista logopédico conlleva a la primordial construcción de un algoritmo de pruebas diagnósticas a través de tecnologías como Inteligencia Artificial, Sistemas de razonamientos basados en caso (CBR), Sistema basado en reglas etc.

El término de dislalia corresponde al trastorno en la articulación de los fonemas, o bien por ausencia o variación de ciertos sonidos concretos o por la sustitución de dichos por otros de manera impropia. Puede dañar a cualquier consonante o vocal, pero frecuente darse una más grande incidencia del problema en determinados sonidos. Las causas de los esfuerzos en determinados sonidos son distintas.

En unos casos podría ser que requieran más velocidad y exactitud de movimientos, como ocurre con la /r/. En otros casos el motivo podría ser que el punto de articulación no es visible y se puede hacer más complejo su imitación, como la /k/. En ocasiones, es ya que hay sonidos, como la /s/, en los cuales existe una más grande tendencia a deformar las posiciones articulatorias de la lengua.

En cuanto a la construcción del algoritmo. Se debe interpretar los datos de los casos obtenidos y para poder construir el algoritmo se requiere de una investigación de campo documental, además de expertos que en conjunto nos ayuden para que estos datos constituyan una base de conocimientos (Regalado Macias & Calle De la A, 2020).

Descripción de la situación problemática

Ubicación del problema en un contexto

En Ecuador en la provincia del Guayas, se obtuvo que Guayaquil tiene una población de 3,645.483 millones, donde el 10% son niños de 5 a 9 años (INEC, 2010). A través de las entrevistas sostenidas con las terapeutas logopedas, según su criterio como profesional es la población infantil la que presenta mayormente el trastorno de la articulación de los fonemas y aclarando que, a su vez, también existe una población adulta con el trastorno mencionado, el cual está acarreado problemas sociales en torno a su círculo familiar, escolar y social.

Una de las principales causas de la dislalia funcional es que no se utilice de forma correcta los órganos articulatorios, el infante con dislalia funcional al momento de articular un fonema no utiliza correctamente estos órganos a pesar de no existir un problema de tipo orgánico. A nivel latinoamericano la carencia de escuelas de fonoaudiología y logopedias es uno de los problemas más recurrentes como lo menciona Marrero (citado en Guillén Escamilla, 2020) “Colombia es el país que tiene más programas (17), seguido por Argentina (16) y Chile (13), mientras que en otros países la presencia es más periférica: Bolivia (2), Perú (2), Ecuador (1)” (pág. 322), considerando esta afirmación es que existe la carencia de profesionales logopédicos y establecimientos médicos que ayuden a contrarrestar estos trastornos del habla en niños, lo cual ocasionará severos problemas a futuro en el transcurrir del ambiente educativo y su vida cotidiana.

Situación conflicto nudos críticos

Los pacientes que presentan este tipo de trastornos dislálicos deben de asistir a varias sesiones de carácter privado y la realización de estos tratamientos conlleva un alto costo, por esta razón no se logran culminar estas sesiones de recuperación sumado a esto la poca atención e importancia que le brindan los padres a este tipo de inconvenientes conllevan a no ser tratados con gran anticipación, finalmente la carencia de profesionales Logopédicos en Ecuador es debido a que existe pocas escuelas fonoaudiológicas y de logopedias siendo este un obstáculo por el cual no se logra avanzar con la recuperación de los pacientes que padecen este tipo de trastornos dislálicos.

A su vez no se cuenta con un proceso que automatice el diagnóstico de pacientes con trastornos dislálicos por lo cual no se ha investigado, planteado ni construido un prototipo de algoritmo que genere diagnósticos en base a la información obtenida de las fichas logopédicas, para así poder iniciar la etapa de terapias y tratamientos adecuados en base a la situación diagnosticada.

Las fichas logopédicas son llenadas de forma manual por los especialistas y en base a anteriores experiencias designan un tratamiento pocas veces efectivo por lo que se busca construir un algoritmo de diagnóstico para que sea considerado a futuro dentro del sistema de razonamiento basado en casos perteneciente al proyecto FCI Herramienta para el tratamiento de alteraciones Fonemáticas (Dislalia) de la Facultad de Ciencias Matemáticas y Físicas, en beneficio directo para los pacientes e indirecto para los especialistas logopedas con el cual puedan optimizar sus procesos mencionados con anterioridad y así poder llevar un historial médico de cómo tratar estos trastornos.

Delimitación del problema

En el campo de la psicología y la logopedia es a donde nos lleva el área de interés, aplicando en la parte tecnológica la construcción de un prototipo de algoritmo para diagnóstico que interactúe directamente entre el especialista Logopeda y el diagnóstico inicial aplicado al paciente, automatizando el algoritmo de esta manera sobre la decisión del diagnóstico, para así conocer si el paciente cuenta o no con dislalia funcional y que déficit de fonemas presenta.

Tabla 1

Delimitación del problema

Campo	PSICOLOGÍA- LOGOPEDIA
Área	Desarrollo de Sistemas
Aspecto	Dislalia
Tema	Algoritmo para diagnóstico de trastorno dislábico en pacientes a partir de la ficha logopédica

Elaboración: Jorge Philco Baque, José Serrano Guamán

Fuente: Anteproyecto

Evaluación del Problema

DELIMITADO

Este proyecto se enfocará en la ciudad de Guayaquil, ya que existe un gran déficit de centros médicos que provean ayuda en la terapia de lenguaje para pacientes con dislalia, es importante recalcar que los pacientes que padecen de estos trastornos también presentan grandes limitaciones en el ámbito del aprendizaje, ya que se les dificulta mucho su entendimiento y a largo plazo si no se logra tratar podrían ocasionar problemas severos a una edad adulta.

EVIDENTE

Con la acción de observar al paciente, se podrá evaluar en los diversos casos su causa física ya sea anomalías en los labios, lengua, nariz y así poder frenar las complicaciones futuras, entre los más comunes se encuentran los pacientes con labio leporino, este impide la correcta formación del rostro, y en la mayoría de los casos este problema es una de las razones por las cuales los pacientes son víctimas de acoso escolar infantil.

RELEVANTE

Sera de gran importancia, enfocándonos en el objetivo primordial de contrarrestar estos problemas de pronunciación y entendimiento en los pacientes que lo padecen, ya sea por la dificultad de comunicación, el cual se debe evitar de manera oportuna para que no provoque mayores anomalías en el ámbito personal.

ORIGINAL

Con la utilización de una herramienta tecnológica que los provea de ayuda a los logopedas a diagnosticar a los pacientes que padecen de estos trastornos y a su vez obtener un tratamiento oportuno según el análisis realizado a cada paciente en el menor tiempo posible y con una alta eficiencia.

FACTIBLE

Con el algoritmo a desempeñar se buscará una pronta solución a dicho problema de dislalia ya que de manera tecnológica se prevé facilitar una rápida y fácil interpretación del mismo, a su vez también se tomará en cuenta los diagnósticos de los logopedas y así poder identificar el tipo de Dislalia que presente cada paciente y en respuesta también conocer cuál sería el tratamiento respectivo que se debe llevar a cabo para el mejoramiento del paciente.

Para que esto sea posible se deberá realizar una recolección exhaustiva de información de las fichas logopédicas, el cual nos permitirá que algoritmo pueda arrojar un diagnóstico y con ello obtener conclusiones incluso si estas pueden ser o no complejas, pues se ha demostrado que en una variedad de casos la Dislalia funcional se manifiesta al inicio, medio y final de un fonema.

IDENTIFICAR LOS PRODUCTOS ESPERADOS

Para poder dar ayuda al proyecto FCI que este a su vez ayudara o beneficiara a los pacientes con trastornos dislálicos y especialistas logopedas se han identificado los siguientes productos: el primer producto esperado es el algoritmo para diagnóstico de los pacientes que a su vez se construye de la investigación de las fichas logopédicas más la investigación de los posibles algoritmos que se usen como base, como segundo producto esperado un dataset simulado construido a partir de datos reales, generando así como último producto esperado el paper o artículo científico.

Causas y consecuencias del problema

En la tabla 2 se indica las principales causas y por ende consecuencias relacionadas al problema.

Tabla 2

Matriz de causas y consecuencias del problema

Causas	Consecuencias
C1. Gran porcentaje de pacientes con trastorno del lenguaje.	E1. Poca disponibilidad de citas médicas.
C2. Falta de especialistas del trastorno del lenguaje	E2. Un sinnúmero de casos sin ser tratados.
C3.1 Desconocimiento de la enfermedad del trastorno del lenguaje.	E3. Con el pasar del tiempo aumentan su porcentaje de gravedad.
C3.2 Poco interés en conocer sobre la enfermedad.	

C4. Falta de algoritmo que sistematice el diagnóstico.	E4. Diagnóstico basado en experiencias pasadas de otros pacientes.
C5. Falta de centros públicos para tratar las enfermedades fonéticas.	E5. Acceso limitado a un diagnóstico y tratamiento temprano.
C6. Valores elevados en el sector privado.	E6. Escasa accesibilidad.
C7. Sistema de diagnóstico de forma manual.	E7. Las consultas conllevan mucho tiempo.
C1. Gran porcentaje de pacientes con trastorno del lenguaje.	E1. Poca disponibilidad de citas médicas.
C2. Falta de especialistas del trastorno del lenguaje	E2. Un sinnúmero de casos sin ser tratados.
C3.1 Desconocimiento de la enfermedad del trastorno del lenguaje.	E3. Con el pasar del tiempo aumentan su porcentaje de gravedad.
C3.2 Poco interés en conocer sobre la enfermedad.	

Nota: Esta tabla refleja las causas y consecuencias que se realizó en base a la recopilación inicial. La elaboración es propia y la fuente de las investigaciones dadas.

Formulación del problema

¿Cuál es el beneficio de aplicar un algoritmo para diagnóstico de trastornos dislálicos en pacientes con problemas del lenguaje ubicados en la ciudad de Guayaquil, como ayuda a los profesionales logopedas que manejan estos trastornos del habla?

Objetivos del proyecto

Objetivo general

Desarrollar un algoritmo basado en reglas para el diagnóstico de trastorno dislálico en pacientes a partir de las fichas logopédicas, usando herramientas de código abierto con la finalidad de identificar la presencia de los respectivos déficits en los fonemas.

Objetivos específicos

1. Realizar una investigación sobre las diferentes pruebas de evaluación denominadas fichas logopédicas, analizando fuentes bibliográficas, artículos científicos, estudio de casos similares para tener argumentos válidos en la construcción del algoritmo.

2. Analizar los diferentes tipos de algoritmo, mediante consultas en trabajos relacionados, antecedentes de estudio y demás fuentes informativas como Google Scholar, bibliotecas virtuales, etc., para determinar cuál de estos algoritmos resulta más eficiente como una solución a la problemática planteada.
3. Construir un prototipo de algoritmo en base a la información de las pruebas de evaluación y criterios de los especialistas logopedas por medio del lenguaje Python, para que pueda otorgar un diagnóstico en base a la información ingresada.
4. Redactar un artículo científico mediante los resultados de la investigación realizada para poder compartir una memoria del trabajo realizada.

Alcance del proyecto

Analizar las fichas logopédicas en las diferentes instituciones y organizaciones de terapias de lenguaje, bases de datos Universidad Guayaquil, en diferentes textos académicos, logopédicos, psicológicos y terapeutas del lenguaje para construir una línea base de factores en problemas con Dislalia.

1. Estudiar las diferentes técnicas para la construcción de algoritmo de diagnóstico en campos logopédicos, médicos o similares en artículos de redacción científica utilizando base de datos indexadas como Google Escolar, Scopus, WOS, etc.
2. Construir el algoritmo en Python, con el editor de código Visual Studio Code.
3. Realizar prototipo donde se introduzcan las entradas de los resultados más destacados de las fichas logopédicas.
4. Establecer planes de pruebas en algoritmo.
5. Redactar el artículo científico.

Justificación e importancia

En la actualidad la tecnología se ve inmersa en todos los ámbitos ya sean estos educativos laborales, científicos, médicos entre otros. Con el uso de aquellas herramientas tecnológicas que hacen posible aprovechar todos sus recursos en su totalidad, es por lo que, en el presente proyecto se trata de adaptar este algoritmo para diagnóstico a que ayude a los centros de salud que tratan las terapias del lenguaje, realizando así un estudio un poco más exhaustivo y minucioso para comprobar que tan viable será el uso de estas herramientas en pacientes con estos trastornos fonéticos.

Se define a la dislalia como una alteración fonética la cual impide que se dé una correcta pronunciación en varios fonemas, teniendo en cuenta que en Ecuador la demanda médica a nivel público de especialistas Logopedas es alta y no se abastece con estos profesionales, mientras que su contraparte en el sector privado existe poca oferta para tratar estos trastornos ya que sus valores económicos son elevados y en la mayoría de los casos los progenitores de los pacientes que padecen los trastornos no cuentan con los recursos necesarios para suplir estas terapias de lenguaje.

La importancia de un prototipo que agilite la toma de decisiones ante un diagnóstico temprano y afirmativo de dislalia funcional permitirá a su vez disminuir este tipo de alteraciones ya que en muchos casos este problema puede presentarse hasta la edad adulta con mayor grado de dificultad y con ello poder erradicar problemas sociales a futuro.

Esto nos lleva a la aplicación de este proyecto, que es de carácter social, educativo y tecnológico el cual mediante la Universidad de Guayaquil pretende resolver un problema inmerso en la comunidad que en este caso son los trastornos de lenguajes, haciendo uso de todos

aquellos conocimientos adquiridos por los estudiantes a lo largo de su ciclo universitario apoyándose con los recursos y herramientas tecnológicas actuales.

Limitaciones del estudio

El presente proyecto se lo realizará en Ecuador en la ciudad de Guayaquil donde actualmente se vive una crisis sanitaria a nivel mundial ocasionada por el COVID-19, el cual nos limita en la realización de actividades de levantamiento de información ya que en varias ocasiones se he decretado de manera obligatoria paralizar todas las actividades en todos los ámbitos, laboral, social, realizando un confinamiento respectivo y que se conserve el distanciamiento entre los habitantes.

Con dicho precedente se hace más complejo la realización de las entrevistas al personal médico especializado en estos trastornos de dislalia lo cual influye en la obtención de información acertada acerca de los procedimientos a realizar para tratar estos trastornos, pero en medida de estas limitaciones se optó por la utilización de la tecnología y dispositivos inteligentes que ayuden a proporcionar información concreta y con ello poder continuar con la elaboración del prototipo.

El tiempo también es otro factor limitante de cierto modo, pues como se supo mencionar la crisis sanitaria que se está viviendo a nivel mundial, influye mucho en cuanto se tomará recolectar la información en las áreas de trabajo donde se tratan este tipo de trastornos, en este caso los accesos restringidos a los hospitales y los pocos profesionales en esta área, como una salida viable se seguirá realizando las investigaciones en artículos científicos, papers y demás fuentes bibliográficas que aporten para cumplir con el objetivo propuesto.

CAPÍTULO II

MARCO TEÓRICO

Antecedentes del estudio

En este apartado se tomara como referencia investigaciones previas de un nivel internacional, nacional y local, tomando en consideración de manera internacional al país de México, investigación realizado para la Universidad Juárez Autónoma de Tabasco, de manera nacional a la ciudad de Cuenca investigación realizada para la Universidad Salesiana y de forma local trabajos investigativos desarrollados en la Universidad de Guayaquil, donde se analizará los resultados obtenidos de aquellos estudios y de ser posible tener referencia sobre el mismo contexto en el que este proyecto estará enfocado.

Investigaciones a nivel Internacional

Alarcón, Marchi, Hernández, Chávez & Hernández Ocaña (2020) manifiestan: en el artículo científico *Clasificadores basados en reglas y selección de atributos para el diagnóstico clínico de subtipos del Síndrome de Guillain-Barré*. Este estudio sirve para analizar el algoritmo basado en reglas, lo cual consiste en utilizar un grupo de normas IF-THEN para la categorización. Una regla IF-THEN es una expresión de la manera IF condition THEN conclusion. La parte IF de una regla se sabe cómo el precedente o precondition de la regla. La parte THEN es la regla consecuente. Estos tienen algunas ventajas, es enormemente expresiva, simple de interpretar y de producir; puede clasificar novedosas instancias y el rendimiento se

compara con árboles de decisiones. También utilizaron algoritmos de aprendizaje automático como por ej.: JRip, OneR y PART, en lo cual obtuvieron resultados a través de los promedios que arrojaron estos algoritmos, los cuales fueron entre 0.4923 a 0.6265, siendo como mayor porcentaje el algoritmo JRip logrando balancear el resultado y obteniendo un porcentaje de confiabilidad reformado.

Este estudio aportó el concepto, análisis y funcionamiento del algoritmo basado en reglas, el cual consiste en utilizar una condición y mostrar un resultado, así sucesivamente hasta completar el o los procesos necesarios y poderla aplicar en la toma de decisiones, obteniendo así una solución más factible para el logopeda.

Investigación a nivel Nacional

Espinoza (2018) indica: en su tesis de grado *Diseño y desarrollo de una herramienta lúdica informática para estimular e intervenir los niveles de lenguaje en niños desde los 4 a 5 años*. La tesis consta de una aplicación multiplataforma, la cual posibilita excitar los diferentes niveles de lenguaje según las necesidades de cada infante, para poder hacer esto se desarrolló un sistema recomendador el cual posibilita examinar los datos del infante y sugerir tal cual la terapia óptima que se debería hacer dentro del programa de estimulación. Se consultaron expertos en la estimulación infantil, especialistas y psicólogos infantiles, para poder conocer bajo que parámetros se debía desarrollar el sistema. Las herramientas que se utilizaron fueron lenguaje Python, minería y análisis de datos aportados por la librería Scikit-learn.

Este antecedente es rescatable para el proyecto dado que se utilizó un algoritmo de clasificación para el análisis de la base de conocimiento, el cual eligió la técnica corpus, es decir, una cantidad infinita de información para poder estructurar los datos, por lo cual esto ayudara a conceptualizar el modelado de la base de datos con las variables adecuadas para la construcción

del algoritmo para diagnóstico. En el algoritmo del árbol de decisión, utilizaron el método de división binario recursivo, para poder dar solución en la estimulación del lenguaje en los niños, esto a su vez ayudo a la categorización del déficit de los fonemas presentes en la dislalia funcional.

Investigación a nivel local

Regalado Macias & Calle De la A (2020) indican: en su tesis de grado *Construcción de algoritmo de similitud para el sistema del proyecto FCI herramienta para el tratamiento de trastornos fonéticos (Dislalia)*. En esta propuesta utilizaron una metodología de prueba experimental para el desarrollo del algoritmo, a su vez mediante entrevistas y encuestas a personas afines al estudio determinaron que el conocimiento en informática es bajo, y que desconocían que la tecnología puede ser usada en beneficio médico, y a su vez el algoritmo de similitud desarrollado ayuda a los especialistas a obtener una información óptima y precisa para pacientes con Dislalia.

Estos proyectos aportaron de manera efectiva para poder realizar el algoritmo de diagnóstico, de manera que pueda ayudar al logopeda a dar un mejor análisis a sus pacientes con resultados eficientes y en el menor tiempo posible.

Fundamentación teórica

Dislalia

La palabra dislalia proviene del griego “dis” que significa “dificultad” y “lalien” que significa “hablar”, se refiere a que es un trastorno o una dificultad para el habla. Es muy usada por logopedas y neurólogos, que denominan habitualmente, a cualquier trastorno del habla, en concreto de la fonación. Según el diccionario de la Real Academia de la Lengua Española, (RAE, 2020) define a la dislalia como “Trastorno del habla originado en el sistema nervioso”.

Otros autores han afirmado lo siguiente:

La dislalia corresponde al trastorno en la articulación de los fonemas, o bien por ausencia o alteración de algunos sonidos concretos o por la sustitución de estos por otros de forma impropia. Se trata de una incapacidad para pronunciar o formar correctamente ciertos fonemas o grupos de fonemas. (Pascual Garcia, 2012, pág. 27)

Para la clasificación de la Dislalia se ha tomado como referencia la investigación realizada por, (Cabrera Cárdenas & Jiménez Cabrera, 2017) en la cual especifican que son las siguientes:

Dislalia Fisiológica o evolutiva

Este tipo de dislalia se presenta cuando el niño/a no articula o distorsiona algunas palabras que escucha, en especial el fonema [r] o ríes, por lo que su desarrollo se debe tanto a una maduración de su cerebro y al aparato fonoarticulador, estas dificultades se superan poco a poco y no deben ser tratadas antes de los cuatro años de edad, si se mantienen más allá de esta edad puede considerarse patológica.

Dislalia Audiógena

Están relacionadas con las dificultades auditivas, el niño o niña que no escucha de forma correcta hablara con defectos, por lo tanto, confundirá fonemas que ofrezcan alguna semejanza al no poseer una adecuada capacidad de diferenciación auditiva. Un especialista debe intervenir para descubrir una dislalia audiógena, y dar un diagnóstico sobre la magnitud de la pérdida auditiva.

Dislalia Orgánica

Aquí tienen la posibilidad de abarcar todos los trastornos de articulación que se encuentren motivados por una causa orgánica. Se denomina disartria cuando el sistema nervioso central (SNC) está afectado, y forman parte de las alteraciones del lenguaje de las personas con deficiencias motoras, por otro lado, se llama disglosia a las malformaciones o anomalías anatómicas de los órganos del lenguaje debido a un defecto de pronunciación, esta deformación suele ubicarse en labios, fosas nasales, maxilares, dientes y lengua.

Dislalia Funcional

Se procederá a analizarla más detalladamente a este tipo de dislalia pues es una parte del enfoque de nuestra investigación.

Pascual Garcia (2012) la define de la siguiente manera:

La dislalia funcional es un defecto en el desarrollo de las articulaciones del lenguaje, por una función anómala de los órganos periféricos, sin que existan trastornos o modificaciones orgánicas en los sujetos sino tan solo una incapacidad funcional. Puede darse en cualquier fonema, pero lo más frecuente es la sustitución, omisión o deformación de r, k, l, s, z, o ch. (pág. 28)

Dislalia Funcional

Déficit fonético y fonológico

Asimismo, dentro de la dislalia funcional, se distinguen dos tipos de trastornos que se integran dentro de esta clase de dislalia, según Alonso (2011, como se citó en López Herrero, 2020) son los siguientes:

- **Trastornos fonológicos:** son alteraciones que se genera a grado perceptivo y organizativo, o sea, en los procesos de discriminación auditiva, perjudicando a los mecanismos de conceptualización de los sonidos y a la interacción entre significante y sentido.
- **Trastornos fonéticos:** son alteraciones de la producción. La complejidad está centrada en el aspecto motriz articulatorio, de forma que no hay confusiones en la percepción y discriminación auditiva. Son infantes con errores estables, o sea, que cometen constantemente el mismo error una vez que emiten el ruido o sonidos problemáticos.

Clasificación según el déficit.

Como específica (Fredes Albarracín, 2019). *Se necesita considerar que la pronunciación de los fonemas puede estar alterada de diferentes maneras.* La clasificación según el déficit en la pronunciación del fonema se da de la siguiente manera:

Sustitución

Error de la articulación en donde un sonido es reemplazado por otro.

Distorsión

Cuando se da de forma incorrecta o deformada.

Omisión

Cuando se omite el fonema que no sabe pronunciar.

Adición

Consiste en intercalar junto al sonido que no se puede articular.

Inversión

Consiste en cambiar el orden de los sonidos.

Factores etiológicos

Los factores en la dislalia funcional etiológicos son:

- Falta de control en la psicomotricidad fina. Se da en los niños o niñas con poca capacidad motora estricta; por falta de ejercicio, mala integración.
- Déficit para reconocer intensidad y timbres entre sonidos, palabras idénticas, fonemas y el poder distinguir diferencias de frecuencias. Lo fundamental es que el niño analiza e integra mal los fonemas correctos que escucha.
- Estimulación lingüística imperfecta.
- También se presentan de forma psicológica: sobreprotección por parte de los padres, temores, celos o poca adaptación, motivo por el cual muchos niños mantienen un lenguaje infantilizado.
- Deficiencia intelectual. Las variedades de dislalias son un trastorno adicional en algunos niños y niñas con deficiencia mental.

Tabla 3

Fenómeno Lingüístico y fonemas alterados

Fenómeno Lingüístico	Fonema Alterado
➤ Sigmatismo	➤ Defecto de la [s]
➤ Kappacismo	➤ Defecto de la [k] [q]
➤ Gammacismo	➤ Defecto de la [g]
➤ Jotacismo	➤ Defecto de la [x]
➤ Lambdacismo	➤ Defecto de la [l]
➤ Rotacismo	➤ Defecto de la [rr]
➤ Deltacismo	➤ Defecto de la [d] [t]
➤ Betacismo	➤ Defecto de la [p] [b]

Nota: Esta tabla refleja los tipos y fonemas alterados respectivamente. La elaboración es propia y la fuente es proporcionada por (Fredes Albarracín, 2019)

Detección de la dislalia funcional

Como mencionan (Cedeño Pinargote & Chinga Quimis, 2020) los posibles pasos para evaluar el nivel de lenguaje de un niño son:

- **Evaluar** de manera general de desarrollo del infante, por medio de observación directa o mediante informes de evaluaciones anteriores.
- **Seleccionar distintos procedimientos para explorar procesos y funciones del lenguaje**, se debe tomar en consideración el nivel de desarrollo del niño/a y sus capacidades perceptivas y motoras.
- **Determinar una secuencia de la evaluación**, con esto nos referimos a realizar varias actividades evaluadoras.
- **Controlar los diferentes tipos de variables dentro de la evaluación**, así como; lugares, horas, estímulos, motivación, memoria, etc.

- **Realizar la evaluación** se debe adjuntar el examen médico general, examen psicológico y evaluación (Sánchez Terán, 2018) del lenguaje.
- **Interpretación de los resultados** el tratamiento consiste en la terapia miofuncional que se debe llevar a cabo entre el odontólogo y el terapeuta del lenguaje.

Tratamiento dislalia funcional

Para el tratamiento de la dislalia funcional se ha tomado como referencia la investigación realizada por, Sánchez Terán (2018) donde especifica que:

Es importante reconocer que se puede aplicar dos tipos de tratamientos, el primero es un tratamiento indirecto el cual está enfocado a superar los problemas de percepción, discriminación auditiva y ejercitación de los músculos fonoarticulatorios y segundo es un tratamiento directo que se enfoca en el problema articulatorio propiamente. (pág. 30)

Tratamiento Indirecto. - el procedimiento indirecto engloba cada una de esas ocupaciones que poseen por objetivo elaborar y facilitar la articulación del lenguaje. Este procedimiento está consolidado por una gigantesca proporción de ejercicios cuya finalidad es lograr la madurez y fortaleza de los órganos fonoarticulatorios.

Para Pascual (2001, como se citó en Sánchez Terán, 2018) las actividades que se deben realizar tienen una correlación con:

- **Discriminación Auditiva** es fundamental tener presente que el infante va a poder pronunciar todos los sonidos de manera correcta cuando haya logrado el reconocimiento y discriminación de estos.
- **Respiración** el lenguaje y el habla poseen una correlación directa, al mejorar la funcionalidad respiratoria y controlarla, los chicos van a poder exponer una mejor articulación de los fonemas.

- **Motricidad** se refiere al desarrollo del desplazamiento vinculado con las capacidades articulatorias como son: velocidad, ritmo, coordinación.
- **Relajación** es oportuno para lograr llegar a una movilidad adecuada de los músculos articulatorios que los infantes logren distencionarlos y controlarlos voluntariamente, debido a la relajación se verán favorecidos los procesos de fonación y articulación.

Tratamiento Directo. - según, Valverde (citado en Sánchez Terán, 2018) afirma que “este tratamiento comprende en la concreción de la articulación correcta de los fonemas que presentan alteración y que sean automatizados de forma general en el lenguaje de los niños” (pág. 32).

Ficha logopédica

Dentro de la exploración de las fichas logopédicas nos encontramos que están conformadas por un conjunto de varias pruebas las cuales son:

Historia Clínica

Consta de datos generales, datos personales, el motivo de consulta , los antecedentes personales, enfermedades en los primeros años de vida del infante, desarrollo psicomotriz físico, como la motricidad gruesa, fina, y de igual manera el desarrollo del lenguaje, antecedentes familiares, antecedentes del niño como vacunas, historia psicosocial, valoración de lenguaje expresivo, exploración de audición a través de sonido, estructura del aparato articulatorio , movilidad del aparato articulatorio y valoración del lenguaje expresivo e impreso.

Ficha de articulación – pronunciación

Consta de una lista de fonemas, cada ficha está compuesta por datos generales del infante y por un grupo de palabras, tomaremos como ejemplo el fonema [p] que puede ir en posición inicial en este caso (pelota), posición media (lampará) y posición final (stop), cada ficha deberá

tener una imagen correspondiente al fonema, todo esto en conjunto nos permitirá obtener información de los problemas que presenta el infante. Como claro ejemplo de una ficha de articulación-pronunciación, se elaboró ELA (Examen Logopédico de Articulación) y posteriormente su revisión, ELA-R (Examen Logopédico de Articulación - revisado).

En el ELA- R Manual de Referencia, ALBOR-COHS (1999) describe:

El ELA-r es un examen completo, en el que se incluye la evaluación del componente fonético-fonológico del lenguaje espontáneo, inducido y reproductivo. Aunque el ELA-R puede resultar una prueba larga, hemos de tener en cuenta que no siempre es necesaria aplicarla en su totalidad. Se recomienda que en todos los casos se evalúe a través del lenguaje espontáneo y el elicitado a través de los dibujos propuestos. En este último, el evaluador decidirá si es necesario aplicarlo al completo o no. Depende de las dificultades concretas de cada niño, de su edad,... entre otros factores. (pág. 14)

Tabla 4

Ficha técnica ELA-r

Nombre	❖ Examen Logopédico de Articulación(revisado)
Autores	❖ Grupo ALBOR-COHS
Administración	❖ Individual
Duración	❖ Variable (15 a 20 minutos)
Niveles de aplicación	❖ 2 años en adelante ❖ Valorar de manera cualitativa el componente fonético-fonológico del lenguaje expresivo en tres modalidades: lenguaje espontáneo, dirigido y repetido
Finalidad	❖ Apreciación del grado de dominio de producción de los diversos fonemas y grupos de fonemas del castellano uso
Significación	

Nota: Esta tabla refleja las características del examen logopédico de la articulación. La elaboración es propia y la fuente es tomada de (ALBOR-COHS, 1999).

Ficha exploración de los órganos articulatorios y de fonación

Esta ficha consta de examen de respiración, exploración de los órganos articulatorios y de fonación, exploración de la motricidad y exploración del estado de la audición.

- **Examen de respiración** consta de respiración nasal, respiración diafragmática, respiración bucal y ejercicios bucales que consisten en coger aire a través de la boca mantenerlo y luego exhalar por la boca, ejercicios de soplo y ejercicios orofaciales como tocarse la nariz con la lengua o morderse el labio superior con los dientes inferiores etc.
- **Exploración de la motricidad** encontramos dos tipos de motricidad, la motricidad gruesa la cual consta de 20 actividades enumeradas, como por ejemplos subir escalones con los dos pies, a su vez para la exploración de la motricidad fina, se cuenta con una tabla de 50 ejercicios, entre ellos realizar bolitas de papel de diferentes tamaños.
- **Estado de la audición** consta de palabras bisílabas, el logopeda debe ubicarse al frente o al costado del niño/a, se dividirá las palabras en listas A, B, C y es necesaria la repetición de las palabras en diferentes metros de distancias, para la lista A seis metros, lista B repetir a tres metros y finalmente la lista C se repetirá mediante un susurro, hasta que el infante no cometa ningún error en la repetición de las palabras.

La prueba EDAF (Evaluación de la Discriminación Auditiva y Fonológica) es una prueba con baremo de fácil aplicación, como se especifica en el sitio web EspacioLogopedico (2021) esta prueba “consta de cinco subtest: Discriminación de sonido del medio; Discriminación figura-fondo; Discriminación fonológica en palabras; Discriminación fonológica en logotomas y Memoria secuencial auditiva” (párr.2).

Tabla 5*Ficha técnica EDAF*

Nombre	❖ Evaluación de la discriminación auditiva y fonológica
Autores	❖ Brancal, M.F.; Ferrer, A.M.; Alcantud, F.; Quiroga, M.
Formato	❖ Manual + Libro de imágenes + CD + 5 protocolos
Duración	❖ 30 minutos
Niveles de aplicación	❖ Desde los 2 años, 9 meses hasta los 7 años, 4 meses.
Edita	❖ Ediciones Lebon, S. L
Destinatarios	❖ Logopedas y Psicólogos

Nota: Esta tabla describe los detalles en general de la prueba EDAF (Evaluación discriminación auditiva y fonológica) para detectar déficits en el ámbito de la discriminación auditiva y fonológica en niños. La elaboración es propia y la fuente es tomada de investigaciones propias.

Exploración del lenguaje expresivo

En cuanto a la exploración del lenguaje expresivo se encontrarán actividades como, nombrar objetos donde el infante indicara lo que visualiza, nombrar acciones donde se le presentara un conjunto de imágenes y debe decir que acción están realizando, a su vez también cuenta con una actividad para nombrar cualidades, donde el infante debe caracterizarlo o describirlo, la exploración del lenguaje expresivo consta de:

- **Exploración de la pronunciación.**
- **Exploración de componente gramatical.**

- **Exploración de la estructura silábica.**
- **Exploración de formas del lenguaje.**

Para la exploración del lenguaje expresivo se cuenta con la prueba psicodiagnóstico ELCE (Exploración del lenguaje comprensivo y expresivo), y según el sitio web EspacioLogopedico (2021) se menciona que “esta prueba aporta una perspectiva global del grado lingüístico del infante destacando los puntos en los cuales se muestra la dificultad, es una herramienta universal para la expresión y comprensión del lenguaje” (párr.1)

Exploración del lenguaje impresivo

La Dra. González González (2005) plantea que “la investigación del lenguaje impresivo incluye el análisis de la audición fonemática, la comprensión de palabras, oraciones simples, estructuras lógico-gramaticales y fragmentos literarios” (pág. 28), esta consta de varias actividades las cuales son:

- Actividades de la comprensión.
- Actividades de oído fonemático.
- Exploración de la lectura y escritura.
- Exploración del ritmo y la fluidez del lenguaje.

Inteligencia artificial

La IA (inteligencia artificial) funciona combinando grandes porciones de datos con procesamiento veloz e iterativo y algoritmos capaces, permitiendo al programa aprender automáticamente de patrones o propiedades en los datos, como Rouhiainen (2018) indica “es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano” (pág. 18).

Esta se lleva a cabo desde algoritmos que son habilidades matemáticas de aprendizaje, y de los datos que hacen falta para realizar estos algoritmos, dichos son datos observables, accesibles públicamente o datos causados en varias organizaciones, los mismos que repiten el proceso para aprender desde ellos.

La limitación primordial de la IA (inteligencia artificial) es que aprende de los datos. No hay otra forma en que se logre integrar el razonamiento. Y cualquier capa adicional de predicción o estudio se tiene que añadir por separado. Aquello supone que cualquier imprecisión en los datos se verá reflejada en los resultados.

Los principales campos de aplicación de la IA han sido en las finanzas, industria pesada, servicios de atención al cliente, transportación, hospitales y medicina. La inteligencia artificial está siendo probada y usada en la industria de la salud para realizar la dosificación de medicamentos, una variedad de tratamientos en pacientes y para métodos quirúrgicos en el quirófano.

Algoritmos

Definición

Un algoritmo es una sucesión lógica y finita de pasos que posibilita resolver un problema o llevar a cabo un objetivo, produciendo un resultado final desde algún punto de partida. En la actualidad los algoritmos están integrados en sistemas y dispositivos electrónicos y se les delegan cada vez más de la toma de decisiones, evaluaciones y analices, que tienen un impacto concreto en nuestras vidas.

La vocación que poseen los algoritmos de penetrar en diferentes zonas de nuestra vida cotidiana ya se ve como un hecho, hacen labores que difícilmente pensaríamos que podrían cumplirse sin un ser humano frente a ellos. Cuando más útiles y sofisticados sean los algoritmos,

más “autónomos” resultan, llegando a dar la impresión de que existe una "máquina de pensar" detrás de algunos de sus enigmáticos razonamientos, una imagen que se remonta a los principios de la informática. De hecho, el concepto de "algoritmo" se frecuenta a ser utilizado o referido como sinónimo de PC, máquina, código, software etc.

A continuación, se han mencionado varios algoritmos, con el propósito de realizar las correspondientes comparaciones entre sí y destacar cada una de las características de estos, comprobando que el algoritmo a utilizar será el apropiado, como Sznajdleder (2017) menciona: “si vamos a diseñar un algoritmo para resolver un determinado problema, debemos tener totalmente estudiado y analizado el contexto de dicho problema” (pág. 3). Conceptualizado de otra manera esto implica, comprender el alcance, identificar los datos de entrada, identificar los datos de salida o resultados.

Una vez dispongamos de un algoritmo que favorezca en gran medida al apogeo y éxito en los resultados esperados con relación al proyecto, y haber realizado las pruebas correspondientes, se explicara cual fue el escogido en la construcción del algoritmo para diagnóstico de trastornos dislálicos, a su vez considerando que funcione correctamente, se necesitan definir criterios para medir su rendimiento o conducta. Estos criterios se centran principalmente en su simplicidad y en el uso eficiente de los recursos.

Clasificación de los algoritmos

Algoritmos Bayesianos

Bayesian Network (BN)

Poseen como fin modelar la dependencia condicional y, por consiguiente, la causalidad, al representar la dependencia condicional por aristas en un gráfico dirigido. Por medio de estas

colaboraciones, uno puede hacer inferencias de forma eficiente sobre las variables aleatorias en el gráfico por medio de la utilización de componentes.

Según Shaier (2019) indica: “these networks can be used for predictions, anomaly detection, diagnostics, automated insight, reasoning, time series prediction and decision making under uncertainty” (párr.13). Estas redes pueden utilizarse para realizar predicciones, detectar anomalías, realizar diagnósticos, obtener información automatizada, razonar, predecir series temporales y tomar decisiones bajo incertidumbre.

Las redes bayesianas tienen la posibilidad de inicializarse de dos formas, dependiendo de si se entiende o no la composición gráfica subyacente:

- 1.- La composición gráfica puede construir un nodo a la vez con distribuciones preinicializadas establecidas para cada nodo.
- 2.- Tanto la composición gráfica como las distribuciones tienen la posibilidad de aprenderse de manera directa de los datos.

Probabilidad

Las BN se usan a menudo para deducir el costo de las variables perdidas dados los valores vigilados. En otros modelos, comúnmente hay un grupo exclusivo o fijo de variables faltantes, como componentes latentes, que tienen que imputarse y, por consiguiente, regresar un vector o matriz fija según tenga sentido las predicciones. No obstante, en la situación de las BN, no tenemos la posibilidad de hacer tales suposiciones, por lo cual una vez que los datos se pasan para la predicción, tienen que estar en el formato de una matriz con NONE en las variables faltantes que tienen que inferirse. La devolución es, por consiguiente, una matriz completa donde los Nones se han reemplazado con los valores imputados.

Inferencia en redes bayesianas

Desde una red Bayesiana ya construida, se da un sistema de inferencia, donde una vez encontradas novedosas pruebas sobre el estado de ciertos nodos, se modifican sus tablas de probabilidad; y dados los valores concretos de algunas variables de una instancia, podrían intentar de estimarse los valores de otras cambiantes de la misma instancia implementando razonamiento probabilístico. El razonamiento probabilístico o propagación de probabilidades se base en esparcir los efectos de la prueba (variables conocidas) a través de la red para conocer la probabilidad a posteriori de las variables.

La inferencia sobre una BN puede presentarse de dos maneras, Soni (2018) afirma:

The first is simply evaluating the joint probability of a particular assignment of values for each variable (or a subset) in the network. The second, more interesting inference task, is to find $P(x|e)$, or, to find the probability of some assignment of a subset of the variables (x) given assignments of other variables (our evidence, e). (párr. 9)

La primera consiste simplemente en evaluar la probabilidad conjunta de una determinada asignación de valores para cada variable (o un subconjunto) de la red. La segunda tarea de inferencia, más interesante, es encontrar $P(x|e)$, o, encontrar la probabilidad de alguna asignación de un subconjunto de las variables (x) dadas las asignaciones de otras variables (nuestra evidencia, e)

Ventajas de las redes bayesianas

Entre las principales ventajas encontramos:

- Posibilitan el hallazgo de la conformación casual implícita en un grupo de datos.
- Permiten disminuir el sobre acoplamiento de los datos.

- Permiten representar unánimemente la dimensión cualitativa y cuantitativa de una dificultad en una interfaz gráfica inteligible.
- Permiten interpretar inferencias bidireccionales; esto es, desde las consecuencias a las causas y desde las causas a las consecuencias, etc.
- Pueden trabajar con datos perdidos de una forma eficaz algo que en la praxis es deseable.

Naive Bayes

Dichos modelos son denominados algoritmos “Naive”, o “Inocentes” en español. Este es un algoritmo de aprendizaje sencillo que usa la regla de Bayes junto con una profunda suposición que las cambiantes predictoras son independientes entre sí. Si bien este supuesto de libertad comúnmente se infringe en la práctica, da una exactitud de categorización competitiva. Junto con muchas de sus propiedades deseables y su eficiencia computacional hacen que Bayes se aplique extensamente en la práctica.

Según indica Shaier (2019) :

A Naive Bayes algorithm assumes that each of the features it uses are conditionally independent of one another given some class. Useful for very large data sets — you can use Naïve Bayes classification algorithm with a small data set but precision and recall will keep very low. (párr.6)

Un algoritmo de Naive Bayes asume que cada una de las características que utiliza son condicionalmente independientes entre sí dada alguna clase. Es útil para conjuntos de datos muy grandes — se puede utilizar el algoritmo de clasificación Naive Bayes con un conjunto de datos pequeño, pero la precisión y la recuperación serán muy bajas. Dado a

que se fundamentan en la libertad de las propiedades con dichos algoritmos se pierde la función de explotar las interrelaciones entre estas.

Puntos débiles y fuertes de Naive Bayes

Los principales puntos débiles son:

- Una vez que el grupo de datos de prueba posee una característica que no fue vista en el grupo de entrenamiento, el modelo le asignará una posibilidad de 0 y va a ser inútil hacer predicciones. Uno de los más importantes procedimientos para evadir esto, es la técnica de suavizado, siendo la estimación de Laplace una de las más reconocidas.
- La suposición de libertad Naive muy posiblemente no reflejara como son los datos en el mundo real.

Los principales puntos fuertes son:

- Inclusive con menos datos de entrenamiento, este algoritmo posee mejor conducta que otros modelos de categorización, en los casos que sea conveniente una suposición de independencia.
- El separamiento de las distribuciones de propiedades condicionales de clase significan que cada repartición podría ser apreciada independientemente como si tuviera una sola magnitud. Esto ayuda con inconvenientes derivados de la dimensionalidad y optimización en el rendimiento.
- Una forma simple e instantánea de pronosticar clases para inconvenientes de categorización binarios y multiclase.

Gaussian Naive Bayes (GNB)

Naive Bayes puede extenderse a los atributos de valor real, comúnmente aceptando una distribución gaussiana. Esta expansión de Naive Bayes se llama Gaussian Naive Bayes (GNB). Estos tienen la posibilidad de usar otras funcionalidades para estimar la distribución de los datos, sin embargo, la distribución gaussiana (o normal) es la más simple de laborar pues únicamente se debe estimar la media y la desviación estándar de los datos de entrenamiento.

GNB admite propiedades y modelos de valor constante, cada uno acorde a una distribución gaussiana (normal). Un enfoque para generar un modelo fácil es aceptar que los datos se describen por medio de una distribución gaussiana sin covarianza (dimensiones independientes) entre magnitudes. Para definir dicha distribución, en los puntos dentro de cada etiqueta, este modelo se ajusta simplemente encontrando la desviación estándar y la media.

Multinomial Naive Bayes (MNB)

El algoritmo Multinomial Naive Bayes (MNB) es un procedimiento de aprendizaje probabilístico que se usa primordialmente en el Procesamiento del Lenguaje Natural (PLN). Calcula la posibilidad de cada etiqueta para una muestra definida y después da como consecuencia la etiqueta con la mayor posibilidad. El algoritmo se fundamenta en el teorema de Bayes y predice la etiqueta de un escrito, como una correspondencia electrónica o un artículo de periódico.

El término Bayes ingenuo multinomial simplemente nos permite saber que cada p ($F|c$) es una distribución multinomial, en lugar de alguna otra distribución. Esto funciona bien para datos que se pueden convertir fácilmente en recuentos, como el recuento de palabras en el texto.

El método MNB usa la frecuencia de términos, es decir, el número de veces que un término definido surge en un documento. La frecuencia de los términos constantemente se normaliza dividiendo la frecuencia de los términos sin procesar por la longitud del archivo. Tras

la normalización, la frecuencia de términos puede utilizarse para calcular estimaciones de máxima verosimilitud basadas en los datos de entrenamiento para estimar la probabilidad condicional.

Algoritmo arboles de decisión

Los algoritmos árboles de decisión, son modelos utilizados para clasificar de forma regresiva los datos; logrando predecir atributos que sean discretos y continuos, cuya principal peculiaridad es el aporte visual que ayuda a tomar decisiones.

Interpretando el punto de vista dicho por Flores (2021), nos indica que este algoritmo “es una estructura jerárquica compuesta por un conjunto de nodos internos y nodos hoja que se encargan de clasificar un conjunto de datos, estos a diferencia de otros modelos son considerados modelos de caja blanca” (pág. 19), por lo tanto, estos permiten que se vea de forma clara la clasificación y se pueda razonar sobre su elección.

Esta misma estructura conocida como árbol de decisión (AD), se presenta como una representación de forma variada lográndose utilizar como métodos de ajustes centrados en árboles de clasificación y llegaron a transformarse en herramientas útiles y sencillas de utilizar. Según Origel, Rendón, Abundez & Alejo (2020), afirman que estos algoritmos trabajan de una forma mixta con datos numéricos o categóricos y estos consisten en crear estructuras que logren clasificar un conjunto de datos. Además, este tipo de algoritmos se basan en las relaciones que tienen los valores denominados como estados con la finalidad de correlacionar y predecir a que clase pertenece el objetivo presentado.

Arboles de decisión

El algoritmo árboles de decisión, es una herramienta que mayormente es utilizada por la comunidad del aprendizaje automático con la finalidad de solucionar ciertos problemas que se presentan en el mundo real con una precisión aceptable.

El funcionamiento de los árboles de decisión y cómo influyen en los diversos algoritmos utilizados es de conocimiento que estos árboles se ajustan a los criterios establecidos para ser aceptados como satisfactorios. La finalidad de estos algoritmos es de convertirse en herramientas de fácil interpretación y que no tengan ningún grado de complejidad en el resultado final. (Pérez Alonso, 2017, pág. 32)

Son considerados modelos descriptivos ya que pueden servir como herramientas explicativas, su función parte del modelo predictivo desde números reales o con valores de clases diferentes, creando un espacio entre el conjunto de diversos modelos de variables representativas. Este tipo de modelo es aplicado para predecir valores numéricos o nominales otorgando una respuesta clasificada, siendo de tipos simples denominados como árboles de clasificación y regresión, definiéndose también como estructuras organizadas de forma jerárquica.

Según León (2018) los árboles de decisión permiten llegar a una resolución final siguiendo una secuencia de condiciones que se logran cumplir desde la raíz del árbol hasta alguna de sus hojas, además, que estos mismos son métodos de Minería de Datos que se emplean para representar los acontecimientos que se presentan al momento de tomar la iniciativa.

Estos modelos son considerados no paramétricos y por lo mismo no tiene que ser necesario que exista una distribución de los supuestos a diferencia de otros modelos clásicos, Quan & Valdez (2018) afirmaron lo siguiente:

A diferencia de otros métodos los árboles de decisión no necesitan la entrada de cualquiera de las divisiones de probabilidad acerca del resultado, además de ser un algoritmo eficaz que puede emplear los datos faltantes; este mismo al encontrar alguna ausencia o anulación de alguna información poco frecuente, aparte de la capacidad de manejar datos puede detectar la no linealidad de efectos y las posibles interacciones entre las variables explicativas. (págs. 2-3)

También se lo define al árbol de decisión como la tendencia de sobre – ajustar (overfit); dichos tienden a aprenderse los datos de entrenamiento muy bien pero no tienen una tan buena generalización.

Las ventajas de utilizar el algoritmo árboles de decisión en un proceso elección son:

- Estos son simples de entender e interpretar.
- Si su esquema no es grande este se visualizará de forma rápida lo cual le da facilidad a que se entienda su desarrollo.
- No solicita que su estructuración de datos sea demasiado exigente.
- Se minimiza la limpieza de datos: los valores faltantes y el outliers no tienen influencia en el modelo.
- No posee restricción para ningún tipo de datos.
- Es considerado un método no paramétrico.
- Se puede utilizar tanto para variables cualitativas como cuantitativas.

Las desventajas del algoritmo árboles de decisión se presentan de la siguiente forma:

- Existe sobreajuste.
- Se pierde información al clasificar variables continuas.

- Su precisión se dice que con métodos como SVM y clasificadores modelo ensamblador de forma recurrente tiene una tasa de error del 30% de forma inferior a CART.
- Es inestable ya que si sucede un cambio en los datos se puede modificar de forma amplia la estructura del árbol y su interpretación no es directa.
- No se garantiza que sea un árbol óptimo.

Random Forest

El random forest o bosque aleatorio son una combinación de árboles que predicen y dependen de los valores de un vector que se maneja de forma aleatoria siendo independiente y distribuyéndose para cada uno de estos. Es también conocido como una técnica de aprendizaje automática que es muy popular teniendo también la capacidad máxima de la generalización para muchos problemas.

Los modelos de predicción como el bosque aleatorio son métodos populares para desarrollar un patrón de predicción en los diversos medios de investigación.

Por otra parte, Lynn et al. (2019) afirman que “el bosque aleatorio es elegido porque a menudo para las predicciones se tiene como objetivo reducir el número de variables necesarias con la finalidad de minimizar la carga de las recopilaciones de los datos de forma eficaz” (págs. 94-95), en consecuencia ante lo expuesto es importante que se evalué diferentes técnicas para al clasificar el random forest este identifique los métodos preferibles y que se base en aplicaciones de sistemas expertos e inteligentes.

Algoritmos de regresión

Los algoritmos de regresión son aquellos que orientan a presentar una prospectiva de información, y debido a esto es necesario que se escoja un algoritmo apropiado que otorgue mejores resultados dependiendo del tipo de datos que se analice.

Según D'Angiolo et al. (2019), partiendo desde la perspectiva de las matemáticas la regresión posee dos significados; el primero surge de la división conjunta de la probabilidad de dos variables aleatorias al mismo tiempo el otro significado es definido como una forma empírica, naciendo así de la necesidad de que sea ajustado entre las funciones de un conjunto de datos. Es interesante entender el punto de vista de esta definición ya que su objetivo es encontrar el prototipo lineal que logre predecir la correlación que existe entre los distintos valores que se deseen encontrar.

A los algoritmos de regresión se los considera como de aprendizaje supervisado y este mismo se establece de forma recta cuando la relación existe entre una variable explicada y una variable explicativa. A este modelo se lo describe también como un subcampo del aprendizaje automático que como objetivo tiene establecer un método para la relación entre un cierto número de características y la variable del objetivo continuo.

Regresión Lineal

Este modelo de regresión es considerado de forma lineal a diferencia de la regresión logística que es utilizada para clasificar datos. La idea que se posee con la regresión lineal es la de optimizar la función de error de la estructura de métodos de mínimos cuadrados, aunque se puede utilizar una ecuación normal para obtener de forma directa la solución de ciertos parámetros.

Según Sandoval (2018), nos menciona que estos prototipos tratan de hallar una línea que se pueda adaptar de una forma excelente a la nube de los puntos dispuestos. También destacan modelos conocidos y utilizados en la regresión lineal esto quiere decir que se acoplan de una forma impresionante a los datos que se poseen; aunque existe el riesgo que nuevos datos no lleguen este es un modelo simple que no ofrecen resultados de manera excelente para interacciones más complicadas.

La regresión lineal es una aproximación donde se configura la relación entre una variable escalar dependiente identificada como “y” y las variables explicativas denominadas “x” en estas podrá haber una o más variables representativas.

Luego de definir la regresión lineal se da paso a una investigación en la cual se puede ver reflejado que estos modelos tienen la desventaja de ser lineales, teniendo en cuenta que las decisiones basadas en la regresión poseen menor probabilidad de que estén sujetas al perjuicio y sean más coherentes, también en este mismo trabajo se afirma que “la línea recta que se forma se ajusta de una mejor manera; identificándose así donde se interpreta “mejor” la diferencia del promedio menor al cuadrado” (Vicuña Pinto, Cortez Chichande , & Basurto Tovar , 2018, pág. 98).

Este tipo de algoritmos son utilizados para definir o cualificar datos que se recolectan y que son provenientes de los diferentes espacios que se desean analizar utilizando estos modelos de regresión. Estos mismos pasan por tipos de datos que pertenecen a diversos indicadores conformándose así conjuntos de variables en el estudio, como antes ya se ha mencionado es una técnica que ayuda a describir variables para obtener respuestas continuas para diversas o solo una variable predictora.

Para (Chen, 2019), el estudio realizado para una comparación que evaluaba la exposición, nos indica que estos algoritmos estadísticos cada vez son más sofisticados y complejos más allá de que esta misma regresión sea ordinaria, sin embargo, es raro que este tipo de algoritmos sean comparados debido a sus diferentes términos en cuanto a su capacidad predictiva.

Regresión Logística

La regresión logística o logistic regression (en inglés), es conocido como un algoritmo de clasificación que es utilizado para predecir la probabilidad de una variable dependiente categórica; es decir que es un modelo logístico binario que es empleado para saber la probabilidad de que una respuesta binaria está basada en una o más variables predictoras e independientes. Este tipo de análisis es utilizado para describir datos, pero a la misma vez logra que se explique la relación entre la variable binaria dependiente y diversas variables o una independiente ordinal, de intervalo o de nivel de razón.

Estas técnicas son utilizadas en diferentes casos como también en diferentes áreas productivas o que trabajan con datos, Sagaró del Campo & Zamora (2019), nos indican en su estudio que “para la identificación de factores pronóstico o de riesgo a partir de diseños observacionales se utiliza la regresión logística para identificar la influencia que tienen diversos factores, aquella es empleada para modelar la cuasi-implicación entre suceso y variables” (pág. 1419), es por lo cual, se comprende que este tipo de algoritmos se emplean para identificar con que influencia se presentan los diferentes factores y lograr solucionar problemas.

Las técnicas son consideradas cuantitativas ya que ayudan a visualizar características del individuo analizado o como estos interaccionan, también reflejan donde estos niveles son agrupados. Merino (2017) afirmó lo siguiente:

En un modelo analítico igual situado en una variable dependiente, aunque para estos de forma inicial sus bases han progresado a variables dependientes con una estructura lineal, también pueden utilizarse en prototipos con variables categóricas y principalmente de tipo dicotómica. El artículo presentado por el autor indica que se debe familiarizar con esta técnica ya que ayudará a proporcionar resultados de forma predictiva. (pág. 179)

Se conoce a la regresión logística como un procedimiento de aprendizaje automático que proviene del campo estadístico, aunque no es un algoritmo para aplicar a los problemas de regresión en la búsqueda de un valor continuo, pero es un sistema para problemas de clasificación. Cabe recalcar que para medir la relación que existe entre la variable dependiente y la afirmación que esta desea predecir siendo dependiente de una o más variables independientes de un conjunto de características disponible para el modelo.

En otro estudio se expone sobre la aplicación de un modelo de regresión logística utilizándola como herramienta de investigación biológicas, entre otras áreas. Chaparro & Cepeda (2018) afirman que su estudio es un método estadístico variado, siendo así empleado en el análisis de datos, aplicándose así también en variables predictoras discretas y continuas. El objetivo era aplicar este modelo de regresión logística para poder determinar factores morfométricos que incidían en las preferencias alimenticias de las diversas aves del bosque de un municipio de Casanare. Logrando así que se pudieran identificar los resultados obteniendo un nuevo modelo de forma eficiente referente a las preferencias alimenticias que se presentaban.

Antes lo mencionado, cabe recalcar que este tipo de método estadístico es automático, que es empleados para clasificar datos, predecir la probabilidad de una variable que es dependiente categórica; permitiendo así que dicha presencia de un factor de riesgo aumente la probabilidad que un resultado se muestre específico.

Algoritmos de agrupamiento

K-Means

K-Means es el algoritmo con el procedimiento de clustering de más fácil uso. Para un enorme conjunto de datos numéricos de alta magnitud se otorga un procedimiento eficiente para clasificar datos semejantes en el mismo clúster (Yu, Chu, Wang, Chan, & Chang, 2018). El algoritmo K-Means se manifiesta vulnerable a los datos ruidosos y valores atípicos, además es susceptible a los centros de cluster iniciales. Yang & Sinaga (2019) indican lo siguiente:

It has been widely studied with various extensions and applied in a variety of substantive areas. Since internet, social network, and big data grow rapidly, multi-view data become more important. For analyzing multi-view data, various multi-view k-means clustering algorithms have been studied. However, most of multi-view k-means clustering algorithms in the literature cannot give feature reduction during clustering procedures. (pág. 114472)

Se ha estudiado ampliamente con varias extensiones y se ha aplicado en una variedad de áreas sustantivas. Desde que Internet, las redes sociales y los big data crecen rápidamente, los datos multi-vistas se vuelven más importantes. Para analizar datos multi-vista, se han estudiado varios algoritmos de clustering k-means multi-vista. Sin embargo, la mayoría de los algoritmos de clustering k-means multi-vista en la literatura no pueden dar una reducción de características durante los procedimientos de clustering.

Racimo

K-Means divide un grupo de N matriz de propiedades de muestra X en conjuntos no interceptados K . Intuitivamente, un conjunto es un conjunto de datos agrupados. Se consideran del mismo tipo los datos en un conjunto y los grupos son el resultado de la agrupación.

Centroide

Comúnmente se llama "centroides" de este conjunto, a la media de todos los datos en un conjunto. En un plano bidimensional, la abscisa y la ordenada del centroide de un conjunto de aspectos de datos, es el costo medio de las abscisas y las ordenadas del conjunto de aspectos de datos.

La labor central de K-Means es hallar K centroides óptimos según nuestro grupo K , respectivamente dedicar los datos más cercanos a dichos centroides y a los conjuntos representados por dichos centroides. En el algoritmo, el número de clústeres K es un hiperparámetro, que necesita nuestro ingreso artificial para decidir.

K-Medians

El objetivo de la agrupación K-Median, al igual que la agrupación KNN, es separar los datos en grupos distintos basados en las diferencias de los datos. Así, al finalizar, el analista tendrá k grupos distintos con características distintivas.

K-medians es un algoritmo no supervisado. Esto quiere decir que el algoritmo tiene la posibilidad de detectar patrones anteriormente desconocidos en los datos sin que los datos tengan que etiquetarse explícitamente.

Entendimiento del algoritmo K-Medians.

- El algoritmo empieza inicializando aleatoriamente los centroides del grupo, los cuales son los vectores localizados en el centro de cada grupo. Estos centroides del grupo se inicializan como aspectos aleatorios en el grupo de datos.
- Después, cada instancia del grupo de datos se asigna a un clúster. Se procede a calcular la distancia de cada instancia, de todos los centroides del conjunto. La instancia se asigna al clúster que tiene el centroide del clúster más cercano. Para la métrica de distancia se utiliza la distancia de Manhattan.
- Una vez asignadas las instancias a un clúster, los centroides del clúster se desplazan. Se desplaza el centroide del conjunto a la postura del vector cuyos recursos son equivalentes al costo mediano de cada magnitud de cada una de las instancias asignadas al conjunto.
- Los clústeres ya se han conformado, no obstante, no hay garantía de que los clústeres particionen bien los datos. Puesto que los centroides del conjunto se inicializaron aleatoriamente. Para superar esto, se procede a reinicializar frecuentemente el algoritmo, y se albergan los mejores centroides de clúster luego de todas las iteraciones.

Algoritmos basados en Instancia

K-Nearest Neighbor (K-NN)

El algoritmo K-NN es la alternativa preferida para predecir la etiqueta de datos de documento sin clasificar basada en atributos de documento. Varias de las causas, Ali et al. (2020) incluyen:

K-NN se discute extensamente en varios tipos de literatura sobre su sencilla utilización, alta eficiencia mejor exactitud y simple de hacer. Comúnmente, ha sido escogido por su simplicidad, y su aplicabilidad con un diminuto número de patrones de entrenamiento. A su vez es el mejor candidato para la categorización basada en "propiedades semánticas". Intuitivamente, es un procedimiento de aprendizaje automático eficaz para solucionar inconvenientes de categorización convencionales. Por estas causas el algoritmo K-NN se estima una de las mejores resoluciones para los inconvenientes.

Basado en reglas

Los sistemas basados en reglas son posiblemente la herramienta más general y conocido de la IA, en inicio está compuesto, por dos piezas principales: un conjunto de reglas que definen su base de reglas (la parte elemental de la base de conocimiento) y el mecanismo apropiado de interpretación de las normas denominado motor de inferencia.

Cualquier sistema basado en reglas está formado por un grupo de normas que definen actividades o conclusiones si se cumplen determinadas condiciones, y un mecanismo de control de inferencia que establece la forma y orden de ejecución de las normas.

La amplitud del proceso de adquisición del conocimiento depende del tipo de conocimiento, como (Reyes González, Martínez Sánchez, Díaz Sardiñas, & Milanés Luque, 2017) indican:

En los Sistemas basados en regla se realiza un desarrollo prolongado y complejo puesto que la sustracción tiene relación con la formalización de normas y el raciocinio humano no constantemente está regido conscientemente por las normas de la lógica; a veces es fundamentalmente un procesamiento de información recuperada con el tiempo.

Ventajas de los Sistemas basados en reglas

Entre las ventajas se pueden destacar las siguientes:

- Tienen una estructura uniforme.
- Personifican de manera natural el razonamiento explícito de los expertos.
- Capacidad para laborar con entendimiento inconcluso e incertidumbre.
- Separación entre la base de entendimiento y su procesamiento.

Análisis de Algoritmos

1. Definición de objetivos e indicadores claves

- **Facilidad de implementación:** Despliegue rápido y de forma fácil a nivel de ejecución o puesta en marcha de una idea programada (algoritmo). (Frascaroli, 2020)
- **Requerimientos técnicos:** Son los puntos técnicos que tienen que cumplirse para terminar exitosamente un plan. Se trata de aspectos como la precisión y fiabilidad. (ERP, 2021).
- **Requerimientos de hardware y software:** Propiedades que debería tener el hardware de una PC para lograr tolerar y/o llevar a cabo una aplicación o tolerar un dispositivo específico. Características que debe tener el software instalado en una computadora para poder soportar y/o ejecutar una aplicación o un dispositivo específicos. (Villabona, 2018).
- **Requerimientos de ambiente:** disponer de herramientas de:
Desarrollo (Sublime, Java, Visual Studio Code, Bases de datos), instalado.
Podría residir en el computador personal del desarrollador.
Tener instalado el mismo manejador de base de datos, versión y calidad.

- **Requerimientos de data: características** de datos estructurados total o parcialmente para ser procesados e interpretados.

2. Identificar los elementos a comparar

En este punto se dejan identificados y definidos brevemente los algoritmos a comparar mediante los indicadores definidos con anterioridad.

- **Naive Bayes.** - Dichos modelos son denominados algoritmos “Naive”, o “Inocentes” en español. En otros términos, que la existencia de una determinada característica en un grupo de datos no está en absoluto relacionada con la existencia de cualquier otra característica. (Roman, 2019).
- **K-NN.**- Es un algoritmo con base en instancia de tipo supervisado de Machine Learning. Puede usarse para clasificar novedosas muestras (valores discretos) o para profetizar (regresión, valores continuos). (aprendemachinelearning, 2018).
- **Basado en reglas.** - En los Sistemas basados en regla se realiza un desarrollo prolongado y complejo puesto que la sustracción tiene relación con la formalización de normas y el raciocinio humano no constantemente está regido conscientemente por las normas de la lógica; a veces es fundamentalmente un procesamiento de información recuperada con el tiempo (Reyes González, Martínez Sánchez, Díaz Sardiñas, & Milanés Luque, 2017).
- **Árbol de decisión.** - Es un algoritmo de categorización y regresión para el modelado de predicción de atributos discretos y consecutivos (Microsoft, 2018).

3. Levantamiento de datos

El levantamiento de datos de los presentes algoritmos descritos con anterioridad, presentaron unas características requeridas para su funcionamiento y ejecución, primeramente, los factores que influyen a nivel de recursos computacionales (hardware y software), sistemas operativos, versiones de S.O, versiones de herramientas de desarrollo, capacidad de memorias rams recomendadas, todo esto que se menciona son de equipos a nivel doméstico para un desarrollo óptimo y pruebas de algoritmos.

Los requerimientos de ambiente se distribuyen, tanto a nivel de herramientas de desarrollo en la cual se precisa los editores de código y herramientas de desarrollo tales como Java, C++, Visual Studio Code, R, Python, los cuales van a permitir la correcta construcción, análisis y comprobación de los algoritmos.

Estudio comparativo herramientas de desarrollo de algoritmos

Como indica el estudio comparativo de Machine Learning: análisis de lenguajes de programación y herramientas para desarrollo realizado para la Universidad Autónoma Baja California. El estudio de datos que se debe llevar a cabo ya sea aprendizaje automático, minería de datos, análisis de datos se seleccionara la herramienta de desarrollo más factible de aprender y poner en práctica (Manrique Rojas, 2020).

Tabla 6

Estudio comparativo herramientas de desarrollo de algoritmos.

Herramientas	Características		
	Velocidad	Curva de aprendizaje	Análisis
Python	Python es dependiente de los paquetes, consiguiente,	es Cuando hablamos de estar orientado a objetos, Python es el lenguaje. Si pertenece al conjunto de	Pero si hablamos de labores en general como el procesamiento de

	vez que hablamos de labores en relación con estadísticas.	programadores funcionales, entonces aprender Python va a ser mucho más simple comparativamente.	datos y el procesamiento de resultados, entonces Python es el lenguaje de programación más apropiado
R	R tiene una virtud comparativamente con Python y es un poco más rápido.	Cuando se trata de la perspectiva funcional, R es el lenguaje de programación	Una vez que hablamos de estudio estadístico, R es el lenguaje de programación apropiado.
Java	Hay mucho que mencionar para aprender Java como un lenguaje de ciencia de datos de primera elección, y además pudimos encontrar que el rendimiento de Java y la estabilidad de tipos resultan muy ventajosos. No obstante, no dispone de una pluralidad de paquetes específicos de estadísticas	Para estudio ad-hoc y aplicaciones estadísticas más dedicadas, la verbosidad de Java provoca que sea una primera alternativa poco factible.	Comparativamente con los lenguajes específicos de dominio como R y Python, Java no dispone de muchas librerías accesibles para procedimientos estadísticos avanzados.

Nota: Esta tabla indica la comparación de lenguajes de desarrollo de algoritmos. La elaboración es propia y la fuente proporcionada por (Manrique Rojas, 2020).

Análisis: Dada la tabla comparativa del análisis de las herramientas de desarrollo, una de las principales variables que destaca en uno u otro sentido son la necesidad de proceso de los datos con los que se va a trabajar. Por ejemplo R es una alternativa recomendada si requerimos hacer un estudio estadístico intensivo cabe destacar que este lenguaje fue desarrollado por y para los estadísticos, python por su lado es más flexible cuando se necesita integrar técnicas de análisis de datos o aplicar procesamiento de datos, sumado a todo esto la simple e inmediata curva de

aprendizaje junto con su versatilidad, realizan de Python un lenguaje de gran calidad para el procesamiento de datos y procesamiento de resultados.

Comparativa de hardware y software

Naive Bayes

Se ejecutó la recopilación de otros trabajos de desarrollo, para levantamiento de los requerimientos de hardware y software para la implementación de los algoritmos, obteniendo como referencia que el autor (Auquilla Morocho, 2021) en su trabajo de titulación denominado “Minería de opinión para textos en español usando procesamiento natural del lenguaje”, utilizo un algoritmo de aprendizaje supervisado Naive Bayes en donde todos los requisitos a usarse se detallan en la tabla 7.

Tabla 7

Características de Hardware y Software

Características	Especificaciones
Ram	8 GB, 2400 MHz
Cpu	4 núcleos, 4 hilos, 2.4 MHz, 6 MB en caché
Disco Duro	500 GB
SO	Windows 10 – 1809 x64
Python	V3.6
Spyder	V5.0
Nltk	V3.6.2

Nota: Requerimientos de Hardware y Software para algoritmo Naive Bayes. La elaboración es propia y la fuente Proporcionada por (Auquilla Morocho, 2021).

Arboles de decisión

También se indago acerca de los requerimientos utilizados en los algoritmos de árboles de decisión, la cual destaca un trabajo de titulación desarrollado por los autores (Garcia Torres & Espino Quiñones, 2018) denominado “Aplicación de minería de datos basado en árboles de decisión para predecir el riesgo de morosidad de los clientes en la empresa de seguros Oncosalud S.A.C. 2018” para poder implementar este algoritmo utilizaron los siguientes requerimientos técnicos como indica la tabla 8.

Tabla 8

Software Disponible

RECURSOS	DESCRIPCIÓN	CARACTERÍSTICAS
HARDWARE	➤ Disco Duro	➤ 1 TB
	➤ Memoria Ram	➤ 8 GB (instalados) / Max 32
	➤ Procesador	GB Tipo DDR3
		➤ Intel(R) Core (TM) i5- 6200
		CPU @2.30 GHz (4 CPUs)
		➤ DATAONE
	➤ Mouse	➤ DATAONE
	➤ Teclado	
	➤ Sistema Operativo	➤ Windows 10, 8 x64
	➤ Microsoft Office	➤ 2016, 2013
SOFTWARE	➤ Netbeans	➤ v8.0, v8.2
	➤ Servidor Web	➤ Apache tomcat v8.5.31 / GlassFish Server v4.1.1
	➤ Xampp / Wampserver	➤ v3.2.2 / v2.4

OTROS

➤ Conexión a Internet

➤ 8MBps

Nota: Requerimientos de Hardware y Software para algoritmo de árboles de decisión. La elaboración es propia y la fuente Proporcionada por (García Torres & Espino Quiñones, 2018).

Basado en reglas

También se obtuvo información acerca de los requerimientos utilizados en los algoritmos basados en reglas, en el cual se hace referencia al trabajo desarrollado por el autor (Martínez Manzo, 2020) denominado “SISTEMA DE MENÚ INTERACTIVO APOYADO POR UN ALGORITMO BASADO EN REGLAS DE ASOCIACIÓN PARA LA ALIMENTACIÓN DE PACIENTES INTERNOS EN UN HOSPITAL. CASO DE ESTUDIO: HOSPITAL DEL PACÍFICO” para poder implementar este algoritmo utilizaron los siguientes requerimientos técnicos como indica la tabla 9.

Tabla 9

Hardware y software de algoritmo basado en reglas

Características	Especificaciones
Ram	8 GB
Procesador	Intel Pentium G3250
Disco Duro	500 GB
SO	Windows 8.1 – x64
Visual Studio	V15.9.4
PostgreSQL	V9.5.18

Nota: Requerimientos de Hardware y Software para algoritmo baso en reglas. La elaboración es propia y la fuente Proporcionada por (Martínez Manzo, 2020).

Acotando que para la implementación de este proyecto se contará con los siguientes requerimientos técnicos tanto a nivel de hardware como de software.

Tabla 10*Características de Hardware y Software*

Características	Especificaciones
Ram	12 GB
Procesador	Intel(R) Core (TM) i7-8550U CPU @ 1.80GHz 2.00 GHz
Disco Duro	1 TB
SO	Windows 10
Visual Studio Code	V1.60.0
Microsoft Excel	V2108

Nota: Recursos de Hardware y Software para implementación de algoritmo baso en reglas. La elaboración es propia y la fuente es propia.

Análisis: Dadas las tablas comparativas del análisis de hardware y Software, una de las principales características que toman en consideración la dimensión de datos a utilizar los tiempos de respuesta a evaluarse y la problemática a resolver. Para los algoritmos de naive bayes se los utiliza en proyectos que requieran algún tipo de entrenamiento previo y que manejen grandes cantidades de datos, por otra parte también se encuentran los árboles de decisiones también manejan gran cantidad de data y se lo utiliza en respuesta a similares resultados dentro de una secuencia de decisiones que se relacionan y de la misma forma se encuentra los algoritmos basados en reglas en los que manejan una mayor información donde su desarrollo se encuentra basado en condiciones, ya sean estos if anidados, o bucles lo cual les permitirá resolver la problemas designada.

Estudio comparativo de precisión de algoritmos

Se midieron varios algoritmos de machine learning como Naive bayes, K-NN y árboles de decisión, para el entrenamiento de estos se usó el dataset construido para el presente proyecto, en este caso el algoritmo naive bayes nos arroja un resultado de 0.13% de precisión como indica la figura , K-NN con un porcentaje de 0.27% , árbol de decisión con un porcentaje de 0.35 % como indica la figura y basado en reglas 0,86% se concluye que para los algoritmos de aprendizaje supervisado, el dataset construido contiene un número elevado de entradas y esto nos arrojaba muchas combinaciones de salida, dado que no existía similitud, estos algoritmos de ML no lograron ser entrenados y por ende realizar la debida predicción del diagnóstico.

Naive bayes

Figura 1

Precisión Naive bayes

```
[45] #Precisión del modelo Naive Bayes
      p2=accuracy_score(y_test, y_pred2)
      print('Precisión del modelo')
      print()
      print(p2)

Precisión del modelo

0.13545816733067728
```

Nota: La grafica indicó que el algoritmo naive bayes arroja una precisión de 0.13 (13%). La elaboración es propia y la fuente proporcionada por las investigaciones dadas.

K-NN

Figura 2

Precisión algoritmo K-NN

```
[40] #Se entrena el modelo
      algoritmo.fit(x_train, y_train)
      #Se obtienen los resultados del entrenamiento
      y_pred = algoritmo.predict(x_test)

[41] from sklearn.metrics import accuracy_score
      accuracy_score(y_test, y_pred)

0.2749003984063745
```

Nota: La grafica indicó que el algoritmo K-NN arroja una precisión de 0.27 (27%). La elaboración es propia y la fuente proporcionada por las investigaciones dadas

Arboles de decisión

Figura 3

Precisión algoritmo árboles de decisión

```
[33] # exactitud del modelo
      from sklearn.metrics import accuracy_score
      accuracy_score(y_test, y_fit)

0.35323383084577115
```

Nota: La grafica indicó que el algoritmo árboles de decisión arroja una precisión de 0.35 (35%). La elaboración es propia y la fuente proporcionada por las investigaciones dadas.

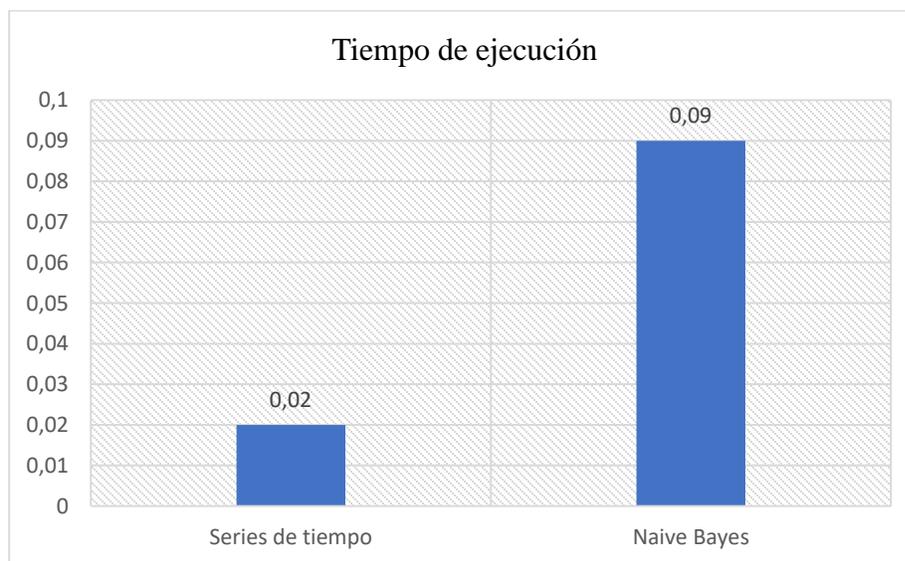
Estudio comparativo de tiempo de ejecución

Naive bayes

Uno de los indicadores que se analizaran y se comparan en esta sección es el tiempo de ejecución de cada uno de estos algoritmos, con lo que respecta a Naive Bayes según investigaciones realizadas el autor (Monja Sandoval, 2020) en su trabajo titulado “Comparación de naive bayes y series de tiempo en la predicción de morosidad de cuotas sociales del colegio de ingenieros del Perú consejo departamental Lambayeque” según sus pruebas con un dataset de 1037 datos obtuvo como resultado que el tiempo de ejecución en Naive Bayes es de 0.09 s.

Figura 4

Tiempo Ejecución Naive Bayes



Nota: Resultado de tiempo de ejecución de algoritmo Naive Bayes. La elaboración es propia y la fuente obtenida de (Monja Sandoval, 2020)

K-NN

Para conocer el tiempo de ejecución del algoritmo K-NN, se realizaron las respectivas investigaciones de la cual (Maillo, Luengo, Garcia, Herrera, & Triguero, 2018) en su respectivo

trabajo titulado “Un enfoque aproximado para acelerar el algoritmo de clasificación Fuzzy K-NN para Big Data” analizaron este indicador comparándolo con otros algoritmos para tener en consideración el tiempo que tardara en procesar toda la data utilizada, arrojando así un resultado como indica la tabla 11.

Tabla 11

Tiempos de Ejecución KNN.

Conjunto de datos	Algoritmo	K	Tiempo total	Precisión
Póker	K-NN IS	3	113,2370	0,4758
		5	123,8311	0,4952
		7	136,5830	0,4937
ECBDL14	K-NN IS	3	27121,0583	0,7833
		5	28673,7015	0,7797
		7	28918,2992	0,7683
Susy	K-NN IS	3	2615,0150	0,6675
		5	2273,6377	0,6784
		7	2372,4100	0,6861
Higgs	K-NN IS	3	10262,9178	0,5454
		5	13446,0519	0,5458
		7	14285,6442	0,5559

Nota: Resultados de tiempo de ejecución de algoritmo KNN. La elaboración es propia y las fuentes obtenidas de (Maillo, Luengo, Garcia, Herrera, & Triguero, 2018)

Arboles de decisión

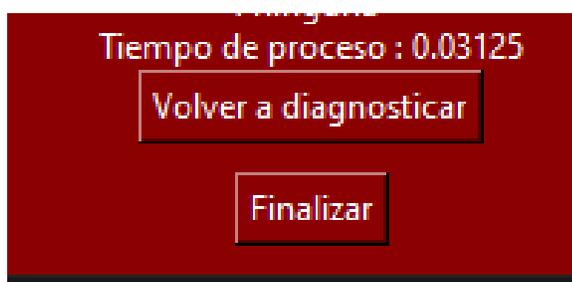
Se realizó el levantamiento del tiempo de ejecución para los algoritmos basados en Árboles de decisiones, del cual el autor (García Girón, 2020) en su trabajo de titulación denominado “Detección de patrones de deserción estudiantil mediante aplicación de Árboles de Decisión C4.5 en el IESTP “Señor de Chocán” de Querecotillo” hizo uso de este tipo de algoritmo, detallando así la cantidad de data utilizada que equivale a un total de 915 instancias respectivamente, arrojaron como resultado un tiempo de ejecución de 0.02 s para su entrenamiento.

Basado en reglas

Finalmente, las pruebas realizadas al presente trabajo investigativo también arrojaron el resultado del tiempo de ejecución que tarda el algoritmo basado en reglas, con una data que contiene 1002 datos respectivamente, los resultados obtenidos son de 0.03125 s, para dar un diagnóstico de los pacientes a evaluarse, siendo este un algoritmo mucho más fácil de implementar se optó por que sea el más adecuado para su desarrollo final.

Figura 5

Tiempo Ejecución basado en reglas



Nota: Resultados de tiempo de ejecución de algoritmo basado en reglas. La elaboración es propia y las fuentes obtenidas de las investigaciones dadas.

4. Tabla comparativa

Tabla 12

Resultados en función de los indicadores vs los resultados de los algoritmos

Algoritmos		Naive Bayes	KNN	Basado en reglas	Árbol de decisión	
INDICADORES	Facilidad implementación	Es de fácil implementación, pero no funciona bien cuando falta cierta combinación de valores en los datos de entrenamiento.	Simple y de fácil implementar, pero tiene un inconveniente importante de volverse significativamente más lento a medida que crece el tamaño de los datos en uso	Implementación fácil, pero se debe conocer de antemano todas las relaciones entre las condiciones los hechos y los datos.	Fácil de implementar y poderoso, pequeñas variaciones en los datos producen grandes cambios en la estructura	
	Requerimientos técnicos en base al dataset	Precisión	0.13	0.27	0.86	0.35
	Requerimientos de software		Windows 8.0, 8.1 y 10 (64 bits)	Windows 8.0, 8.1 y 10 (64 bits)	Windows 8.0, 8.1 y 10 (64 bits)	Windows 8.0, 8.1 y 10 (64 bits)
	Requerimientos de ambiente	Herramienta de desarrollo	Python	Java	Visual Studio Code	Java
			Java	Visual SC	C#	R
		C#	Python	Python	VSC	
		VSC	R	Python	R	
Manejador dataset	MySQL	MySQL	MySQL	MSQL Server		
Requerimientos de hardware		Procesador 2.4 GHz o superior	Procesador 2.3 GHz o superior	Procesador 2.0 GHz o más rápido	Procesador 2.3 GHz o superior	
	Requerimientos de hardware	6 Gb de RAM 8 Gb recomend	6 Gb de RAM 12 Gb	6 Gb de RAM 8 Gb	6 Gb de RAM 8 Gb	

		ado	recomendado	recomendado	recomendado
		Tarjeta de vídeo que admita una resolución de pantalla mínima de 720p (1280 x 720)	Tarjeta de vídeo que admita una resolución de pantalla mínima de 720p (1280 x 720)	Tarjeta de vídeo que admita una resolución de pantalla mínima de 720p (1280 x 720)	Tarjeta de vídeo que admita una resolución de pantalla mínima de 720p (1280 x 720)
	Requerimientos de data en base al dataset de este proyecto.	1800 datos para entrenarse	1500 datos para entrenarse	No necesita entrenarse	2000 datos para entrenarse
	Mayor asertividad			x	
	Mayor requerimiento		x		
	Menor requerimiento			x	
	Residir en el computador del desarrollador o en la nube	X	x	x	x

Nota: La gráfica indica los indicadores cruzados con los resultados de los algoritmos. La elaboración es propia y las fuentes obtenidas de las investigaciones dadas.

5. Interpretación de resultados y conclusión.

Inicialmente no parecía muy preocupante el comportamiento que iban a tener los algoritmos de aprendizaje supervisado utilizando un número tan elevado de entradas sin embargo se interpreta que, en base a los indicadores aplicados en la tabla de comparación, los algoritmos de aprendizaje supervisados han presentado porcentajes bajos en la precisión y unos parámetros muy similares en cuanto a requerimientos de data (1000 datos en adelante) en base al dataset construido para este proyecto, por el contrario el algoritmo basado en condiciones no necesita de un entrenamiento para procesar el dataset.

De la siguiente manera se puede concluir que los mejores resultados, con unos valores muy buenos los ha proporcionado el basado en regla, en este último la buena precisión obtenida sobresale a pesar de tratarse de un algoritmo bastante sencillo queda claro que a pesar de su simple idea su potencia y precisión son elevados, el resultado menos satisfactorio se ha obtenido del algoritmo naive bayes, ahora bien los recursos computacionales requeridos tanto a nivel de hardware como software es evidente las diferencias y similitudes entre cada uno de estos, debido a esto se ratifica la decisión de utilizar el algoritmo basado en reglas para el desarrollo del presente proyecto, pues no requiere de muchos recursos para su aplicación con la finalidad de obtener un diagnóstico preciso al momento de la evaluación de los pacientes.

Herramientas de desarrollo de software

Python 3.9

Python es un lenguaje de programación que se destaca por su código limpio y legible, además de contar con licencia de código abierto por lo que se puede manejar en cualquier escenario. “Es un gran lenguaje de programación interactivo, interpretado y orientado a objetos, combina una potencia notable con una sintaxis muy clara” (Python, 2020).

Algunas de las características de Python que se pueden mencionar, es que es un lenguaje multiplataforma, tanto para Mac, Windows y Linux y es muy ventajoso para ejecutar su código fuente, por lo que se adapta sin muchos cambios, su lenguaje es interpretado, usa tipado dinámico, fuertemente tipado. Soporta varios paradigmas de programación como, estructurada, orientada a objetos, y en menor medida programación funcional. La indentación (formato del código) en Python se da forma estructural.

Comparación de Python frente a otros lenguajes

El presente proyecto se desarrollará con el lenguaje de programación Python, por ser adaptable, simple, flexible y claro, aunque simple es un lenguaje poderoso que puede ser usado para construir grandes proyectos, además de contar con una amplia gama de librerías para la realización de modelos predictivos que ayuden en la generación de diagnósticos. Una vez definidas las características del lenguaje Python y conociéndose su definición, se elaborada una tabla comparativa frente a los otros lenguajes, en este caso el lenguaje R.

R es un Lenguaje de programación compuesto por un grupo de herramientas bastante flexibles que tienen la posibilidad de ampliarse de forma sencilla por medio de paquetes, librerías o explicando nuestras propias funciones elaborado para su uso en estudio estadístico. Su filosofía open source, posibilita que cada una de los instrumentos y funcionalidades hechas por los usuarios logren ser compartidas de forma gratuita y se logre entrar hasta la última línea de código para consultarla.

Tabla 13

Comparación entre Python y R

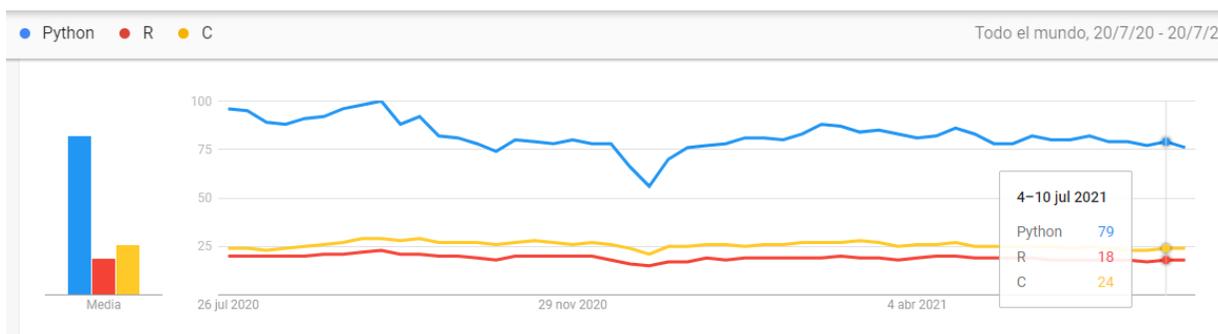
Python	R
❖ Lenguaje de alto nivel multipropósito	❖ Potente en visualización de información y datos.
❖ Lenguaje Open Source	❖ Su instalación cuenta con funciones bastantes completas
❖ Cuenta diversidad de posibilidades y funcionalidades, tiene una sintaxis manejable y simple.	❖ R es un lenguaje más lento en ejecución
❖ Librerías como Pandas Numpy, que implementan funciones para cálculos estadísticos y matemáticos.	❖ Lenguaje muy actualizado referente a técnicas estadísticas.

Nota: Esta tabla compara característica de dos de los lenguajes más populares. La elaboración es propia, la fuente es proporcionada por (Fuentes & Medina, 2021)

En la figura 6 se muestra que el lenguaje de programación Python cuenta con el 79%, frente a C con un 24% y por último el lenguaje R con un 18%. En este caso Python posicionándose en primer lugar en cuanto a herramienta más utilizada durante los últimos 12 meses.

Figura 6 Línea de tendencia comparativa entre, Python, R y C

Línea de tendencia comparativa entre, Python, R y C



Nota: Se visualiza la siguiente línea de tendencia, de los lenguajes más utilizados entre julio del 2020 y julio del 2021. La elaboración es propia, la fuente es proporcionada por Google Trends.

STAT::FIT

Su desempeño es intuitivo, sin embargo, su documento de ayuda es largo. La funcionalidad `Auto::Fit` se adapta automáticamente a distribuciones sucesivas, otorga comparaciones relativas entre tipos de repartición y una medida absoluta de la aceptabilidad de cada repartición. Las propiedades integran estadísticas descriptivas, estimaciones de límites, pruebas de bondad de ajuste, estudio gráfico, generación de cambiantes aleatorias, exportación a productos de simulación (Geer Mountain Software Corp, 2016).

Este instrumento es una de las posibilidades que ofrece el programa de ProModel, el cual es eficaz para hacer la respectiva simulación de datos, se lo usó con el fin de obtener las

distribuciones probabilísticas de cada una de las variables del dataset con la finalidad de posteriormente hacer la Simulación de Montecarlo.

@Risk 8.2

Es un instrumento eficaz para la utilización en Excel, sirve para la construcción de Simulación de Montecarlo. Se lo ha empleado para la obtención de una más grande proporción de datos, un total de 1002 datos, para más adelante añadirlos al dataset.

SPSS

Fue eficaz para hacer estadísticas descriptivas, el cual arroja los próximos resultados: el estudio de correlación Pearson, el chi-cuadrado y las tablas de contingencia.

Revisiones sistemáticas

Al momento de hacer este proyecto se formuló la siguiente pregunta: ¿Cómo generar el diagnóstico a los pacientes con trastorno dislálicos? Se investigó por medio de la aplicación Publish or Perish (Software que cita fuentes de datos diversos incluido Google Scholar y Microsoft Academic) en base a las palabras claves que son: Algoritmo de diagnóstico, dislalia y logopeda.

Los criterios de inclusión y exclusión aplicados fueron los siguientes: el trabajo esté enmarcado entre los últimos 5 años, que tenga pertinencia con el algoritmo de diagnóstico con pacientes dislálicos, las pruebas a evaluar para las personas que sufren de este trastorno, cualidades o características en común para realizar los análisis respectivos, las herramientas que se utilizaron para darle solución al problema.

Metaanálisis

El metaanálisis es una poderosa herramienta para hacernos una idea de lo que a veces se llama el “estado del arte”, es una metodología para el análisis cuantitativo que presenta características como precisión, replicabilidad y objetividad. Permite hacer esta valoración, lo que implica obtener estimaciones del tamaño del efecto, y analizar los matices con que se muestra dicho efecto. (Botella & Zamora, 2017)

Diseño del metaanálisis

En el diseño del metaanálisis se tomaron en consideración la cantidad total de 300 artículos científicos. De los cuales 114 artículos formaron parte de la muestra que cumplieron el 50% de las palabras claves definidas en la investigación bajo el criterio de inclusión y exclusión.

Tabla 14

Fuentes bibliográficas.

N°	Bibliografía
1	Google académico
2	Elsevier
3	Knowledge Center
4	DSpace
5	Springer

Nota: La tabla muestra las fuentes bibliográficas y bases de datos utilizadas para la búsqueda de los artículos científicos que forman parte del metaanálisis. La elaboración es propia y la fuente es proporcionada por las investigaciones correspondientes.

Instrumentos utilizados para la recolección de datos

Se investigó mediante la herramienta “Publish or Perish”, la cual es un software de búsqueda y análisis de artículos científicos. Los parámetros de búsqueda se efectuaron utilizando las palabras claves establecidas en el proyecto, encontrando diversos artículos que cumplan con

las coincidencias de búsqueda establecidas e identificando diversas fuentes bibliográficas las cuales se fueron analizando, para realizar una hoja de cálculo que se procedió a examinar posteriormente, la cual está constituida por autores, título, año, publicación y artículo URL, tal cómo se muestra en el anexo 8.

Pregunta científica a contestarse

Con el presente trabajo de investigación se busca ayudar al logopeda a obtener un mejor diagnóstico acerca de sus pacientes que sufren de trastorno dislábico, por ende, brindarle un mejor tratamiento.

Pregunta:

¿Será posible construir a partir de las fichas logopédicas ELA y PAF un algoritmo que diagnostique los trastornos dislálicos en pacientes?

Variables de la investigación

Las variables tomadas en cuenta para el presente trabajo investigativo se fueron determinando en el transcurso de las revisiones de los artículos y papers relacionados al tema propuesto, y con ello se determinaron las siguientes:

Variable Independiente: Porcentaje de eficiencia del algoritmo.

Variable dependiente: Algoritmo para el diagnóstico del trastorno dislábico.

Definiciones conceptuales

Aprendizaje supervisado

Una PC que distingue y reconoce los conceptos de imágenes u objetos mediante ejemplos anteriormente proporcionados por el hombre.

Diagnóstico

En específico, es un término que se conforma por el prefijo diag- que significa “a través de”; el término gnosis que es un sinónimo de “conocimiento”, y al final el sufijo –tico que se define como “relativo a”. (Rivas, 2021).

Excel

Excel es un programa informático de hojas de cálculo electrónicas realizado por la compañía Microsoft y viene del vocablo en inglés que tiene relación con algo o alguien que "predomina entre los demás" o “es mejor que los otros”. (Significados, 2017)

Fonema

En lingüística (y más específicamente, en fonología) se sabe cómo fonema a la unidad mínima del ruido de una lengua, mejor dicho, a una abstracción teórica que se corresponde con todos los sonidos articulados en el momento de dialogar un lenguaje. (Raffino, 2020).

Inteligencia Artificial

Es el campo científico de la informática que se reúne en la construcción de programas y mecanismos que tienen la posibilidad de demostrar comportamientos considerados capaces. En otros términos, la inteligencia artificial es el término de acuerdo con el cual “las máquinas consideran como seres humanos”. (Salesforce Latinoamérica, 2017).

Logopeda

El logopeda se ocupa de evaluar, diagnosticar, profetizar, rehabilitar y prevenir los trastornos de la comunicación humana. (Logopedia, 2019).

Machine Learning

Shalev-Shwartz & Ben-David (citados en Maisueche, 2019) indica que es una técnica vinculada a la detección automática de patrones importantes en un grupo de datos. En los últimos años, se convirtió en un instrumento bastante común en básicamente cada una de las labores que necesitan sustraer información desde enormes cantidades de datos.

Python

Lenguaje de programación dinámico y sencillo de aprender, posee una sintaxis que ayuda a hacer un código legible cuenta con una multitud de librerías disponibles y un nivel de abstracción alto, lo que consigue que sentencias que en otro lenguaje son largas se puedan realizar en pocas líneas de código (Caturla & Congosto, 2016).

Trastornos

Se habla de una variación conjunta de las condiciones aparentemente clásicos de un organismo. (Pérez, 2021).

Visual Studio Code

Es un editor de código fuente que posibilita laborar con diferentes idiomas de programación, es gratuito, de código abierto y nos da una utilidad para bajar y gestionar extensiones con las que tenemos la posibilidad de personalizar y potenciar este instrumento. (Microsoft, 2021)

CAPÍTULO III

METODOLOGÍA DE LA INVESTIGACIÓN

La metodología de la investigación, según Hernández Escobar et al. (2018) indican: “se ocupa del estudio de aquellas cuestiones que posibilitan la aplicación del método científico de manera rigurosa garantizando que las conclusiones de una investigación se obtengan realmente, a partir de los datos obtenidos, es decir que tengan carácter objetivo” (pág. 29).

En el presente capítulo se dará a conocer la aplicación de la metodología a usarse, las técnicas de recolección de datos, la segmentación de la población y muestra que abordaremos, el tipo de investigación aplicado, la modalidad de la investigación e identificaremos a los beneficiarios o involucrados de forma directa e indirecta dentro del proyecto.

Modalidad de la investigación

Según indica Galeano (2020) “se concibe como forma de investigar en las que se consolida de manera estructurada una opción epistemológica, un propósito, una ruta metodológica y un arsenal instrumental” (pág. 55). El presente proyecto de investigación se ejecuta con una modalidad de campo bibliográfica. En la indagación de campo se ha realizado una entrevista para decidir las variables de análisis establecidas. Por otro lado, del total de la indagación, se hizo alrededor de un 70% de revisiones bibliográficas de fuentes electrónicas, con el propósito de sustraer información importante que sustente la ejecución de la investigación y el proyecto y el 30% de la modalidad es de campo, debido a que se hará un entrenamiento al

algoritmo basado en reglas que consiste en realizar una condición y presentar los procesos sucesivamente hasta terminar con los diferentes diagnósticos dados para cada paciente.

Tipo de investigación

El tipo de investigación utilizado se encuentra basado conforme a la intención, en aquel se encuentra una investigación descriptiva, donde los valores que se emplean para dicha investigación es lograr el diagnóstico sobre el trastorno de los fonemas que presentan los evaluados; en base a los datos otorgados de la base de datos real y contruidos en un nuevo dataset en la cual se desea obtener un diagnóstico sobre las dificultades que presenta el individuo.

Diseño metodológico de la investigación

Determina la estructura de los procesos a realizarse en una investigación. Constituye la especificación y descripción de todos los métodos, procedimientos y tipos de análisis, e artefactos de recolección de datos usados, para comprobar la hipótesis (Solíz Plata, 2019) . Y así llevarla a cabo satisfactoriamente, explicando de qué forma se tomarán, examinarán los datos y qué tipo de pruebas se realizarán.

El diseño de corte transversal se cataloga como un análisis observacional de base personal que frecuente tener un doble objetivo: detallado y analítico. Además, es conocido como análisis de prevalencia o encuesta transversal; su objetivo fundamental es detectar la frecuencia de una condición o patología en la población estudiada y pertenece a los diseños básicos en epidemiología al igual que el diseño de casos y controles y el de cohortes (estudio al que se observa a lo largo de un lapso). (Rodríguez & Mendivelso, 2018).

Se utilizó este diseño porque se analizará dos variables, que son cuantitativas y cualitativas:

Cuantitativas: Son los fonemas que forman el dataset y están determinados en el rango 0 - 1.1 - 1.2 - 1.3 - 1.4 - 2.1 - 2.1 - 2.3 - 3.1 - 3.2 - 3.3 - 4.1 - 4.2 - 4.3 - 5.1 - 5.2 - 5.3

Cualitativas: Al momento de ingresar al diagnóstico el dataset se convierte en variables no medibles para poder trabajar con dicha información.

La presente indagación se desarrolló con un modelo de campo, se desarrollaron instrumentos recolección de datos (entrevista) y técnicas para decidir las variables de análisis. Del mismo modo para la recolección de información relacionada con la investigación se procedió a una revisión bibliográfica de artículos científicos (investigación documental) y bases de datos académicas.

Metodología de investigación

Investigación descriptiva

Según indican Burgo, León, Cáceres, Pérez, & Espinoza (2019)

Las investigaciones descriptivas, son aquellas que observan y describen el fenómeno; no requieren la manipulación de variables. Se desarrollan para conocer las características más relevantes del estado del fenómeno y los factores asociados al problema. En muchas ocasiones tienen función diagnóstica. (pág. 322)

Etapas del método descriptivo

El método descriptivo es denominado de carácter cualitativo y se utiliza en investigaciones que tiene como objetivo que ayude a valorar ciertas características de una situación en particular u objeto en estudio.

1.- Identificación y delimitación del problema

Es considerado el primer paso del estudio, tratándose así del punto exacto donde se decide que se investigará y como se formularan las preguntas para obtener respuestas. Para tener un estudio de forma correcta se debe de identificar el problema de la investigación, debido a que es de forma necesaria lograr determinar aspectos centrales que como todo problema posee al encontrarlo (Schwarz Díaz, 2018); el propósito de este campo de estudio es comprender de mejor forma el aspecto para realizar una determinación en ciertos espacios inspiradores.

2.- Elaboración y construcción de los instrumentos

Esta etapa del proceso debe de efectuarse con cierta antelación, ya que es de importancia que se asegure la información que se desea obtener. Además, es un componente determinante para quien realiza el estudio; para de forma legítima cuantificar la información para la que fue diseñado con la finalidad de conocer puede presentarse en fase cualitativa que es la elaboración de un contenido o se presentan ciertas preguntas o ítems (Martínez R., 2019). Por tal motivo la elaboración y construcción de un instrumento debe de tener claridad de conceptos sobre estructuras teóricas, medición, confiabilidad y validez.

3.- Observación y registro de datos

Es aquella fase del método descriptivo que ayuda recolectar información visual sobre la situación que se desarrolla; además que se estudia como el objeto en cuestión se comporta. Cabe recalcar que esta parte del proceso es crucial ya que implica estar conectado con la realidad observada para así tener en cuenta la mayor cantidad de la forma más detallada posible asimismo es recomendable que este tipo de observación no altere condiciones naturales en las que se presenta el fenómeno estudiado (Romero Tabeayo & Rosa N., 2017). Es importante mencionar que es fundamenta en una investigación de carácter cualitativo, proponer diversos modelos de registros para la observación en la fase de campo.

4.- Decodificación y categorización de la información

En esta etapa se obtiene los datos y se transcriben en algún formato, por lo cual la finalidad de esta fase es la de proceder a interpretar un estudio con la intención de que este tipo de trabajos de investigación generen conocimientos profundos y sistemáticos del fenómeno estudiado (Montoya Zuluaga, 2020). En consecuencia, esta fase permite que los datos sean procesados con mayor facilidad cuando se trata de cantidades grandes o de diversas categorías que pueden confundirse.

5.- Análisis

Esta parte del proceso es la cual una vez que los datos han sido catalogados categorizados, en esta fase será el momento de interpretar y analizar mediante referencia la investigación en cuestión; por tal motivo es importante señalar que en el eslabón antes mencionado con la información recabada en su análisis no se debe establecer relaciones causales ya que su metodología no lo permite (Montoya Zuluaga, 2020).

6.- Propuestas

Aquel es el momento del proceso donde se toman sugerencias de pasos a seguir para investigar el fenómeno u objeto de estudio, con lo antes analizado es normal que aparezcan preguntas y es aquí donde se propone que se indague con mayor exhaustividad ese tipo de cuestiones (Cobo Viera & Jaramillo Armijos, 2020).

Población y muestra

Población

Se hace referencia a un conjunto que posee todos los componentes o individuos que poseen una misma naturaleza, evidenciando así una cualidad propia o que, en su defecto, se lo define con un mismo concepto y que cuyos componentes se los somete a un análisis que es

referente a sus características asociándose de diversas formas. Se puede denominar también a la población como un grupo de unidades de diversas índoles que se desean estudiar (Otzen & Manterola, 2017).

Para el desarrollo del siguiente proyecto se tomó en cuenta un conjunto de población como se muestra en la tabla 15, dentro de la ciudad de Guayaquil, el cual esta conforma por 10 especialistas en el trastorno de la dislalia denominados logopedas de diferentes hospitales e instituciones, los cuales nos colaboren mediante una entrevista para conocer información referente a los criterios que se tomen en cuenta al momento del diagnóstico, que fichas logopédicas más utilizan, tiempos y criterios del diagnóstico , debido a la pandemia que se vive actualmente las restricciones que esta conlleva y al no poder realizar entrevistas de manera presencial se optó por un número no muy elevado de especialistas a los cuales se les aplicara una entrevista online, toda la información recolectada servirá para la construcción y culminación del algoritmo.

Tabla 15

Hospitales o Instituciones de especialistas del lenguaje

Población	N°
Hospital de Niños Dr. Roberto Gilbert	5
Instituto de Neurociencias	5
Total	<u>10</u>

Nota: La tabla muestra la población objetivo del proyecto. La elaboración es propia y las fuentes de las investigaciones realizadas.

Población Objetivo

Se la define de aquella forma al grupo de personas que son de interés para realizar un estudio, la denominada población objetivo es la que se puede ver o termina siendo afectado por la investigación que se realice, también se le denomina de esta manera porque dichos individuos serán beneficiados con el proyecto y estos se pueden identificar partiendo de un individuo o elemento de referencia que es definido como el total del área (Barboza & otros, 2020). Dicho esto, para la presente investigación se utilizará una población de niños entre una edad definida para que sus fichas de evolución sean evaluadas por el algoritmo desarrollado.

Muestra

Es una parte o división del universo o la misma población en donde se lleva a cabo una investigación en la cual existen procedimientos para obtener la cantidad de elementos que posee la muestra como lógica, formulas u otros que con el paso de las etapas se observaran más adelante. Definiendo lo que es una muestra en si se puede decir que es una cantidad pequeña o un subconjunto de una cosa la cual es considerada representativa de su totalidad y que se puede tomar o separar de ella por diversos métodos para someterla a un análisis o experimentación (Robles Pastor, 2019).

Introducción

Para obtener con mayor exactitud la cantidad de profesionales logopedas actuales en Ecuador se indago y consultamos en el siguiente sitio web www.espaciologopedico.com, al cual nos dirigimos mediante correos electrónicos solicitando que nos ayudaran con la información mencionada para poder continuar con el presente trabajo investigativo, lastimosamente no obtuvimos una respuesta positiva por parte del sitio web, de igual manera del Hospital de Niños Dr. Roberto Gilbert y el Instituto de Neurociencias no se obtuvo respuesta por ende se tuvo que

buscar otras fuentes por lo cual solo se pudo contactar con 2 profesionales logopedas donde se les aplico una entrevista respectiva para que nos proporcionen toda la información necesaria que aporte al presente proyecto.

La muestra tomada para este proyecto fue seleccionada mediante la técnica del muestreo aleatorio. La cual consiste en tomar una pequeña muestra de una población ya asignada para su posterior análisis y esta detallada a continuación:

Muestra

- ❖ 2 terapistas del lenguaje o también denominadas logopedas.
- ❖ 3 profesores de distintas áreas de conocimiento en la informática.

Tabla 16

Población de la ciudad de Guayaquil para entrevista.

POBLACIÓN DE LA CIUDAD DE GUAYAQUIL	N
Especialistas logopedas	2
Profesores de la Facultad de Ciencias Matemáticas y Físicas	3
TOTAL	5

Nota: La tabla muestra el tamaño de la muestra tomada mediante la técnica del muestreo aleatorio. La elaboración es propia y la fuente de las investigaciones realizadas.

Muestreo

Esta definición corresponde a una acción que analiza las asociaciones que deben de existir al momento de una distribución de variables “y” en otra población denominada “z”, distribuyendo así a esta con la muestra del estudio. La estadística define a la muestra como una división de una población determinada en la cual puede haber dos tipos de muestreo que son los fundamentales como el aleatorio o probabilístico y el no aleatorio que también se considera no probabilístico. También se puede denominar un proceso o un grupo de métodos que ayudaran a

obtener una muestra finita de una población o conjunto de individuos finita o infinita (Hernández & Carpio, 2019).

Marco Muestral

El marco muestral es considerado un conjunto de unidades que intervienen en un análisis de forma delimitada por ciertos parámetros muestrales, ayuda a seleccionar asegurando que la probabilidad de que todas estas variables sean escogidas para esta uniformidad existe los marcos de listas donde los antes mencionados pueden ser enumerados uno por uno. Se puede definir que es una lista de elementos que forman el universo que se estudiará y que cada uno de estos elementos tiene características que les permiten lograr identificarse (Guzmán Rizo, 2020).

Procesamiento y análisis

Técnicas de recolección de datos.

Según Hernández (citado en Galvez Quilca, 2019) indica que “la implementación de las técnicas de recolección de datos contempla elaborar y realizar un procedimiento bien ordenado que posibilite recopilar estos datos con un objetivo definido” (pág. 64). Para garantizar la calidad de dichos datos la organización debería conceptualizar las fuentes donde se obtendrá la información, la ubicación donde se encuentra la fuente de información, el proceso en el que se recopilaran estos datos y el método y análisis a utilizar con el fin de responder a la problemática planteada. Para el presente proyecto se utilizó la técnica de la entrevista y encuesta, la cual permitió recolectar directamente de la muestra seleccionada los datos primarios.

Instrumentos de recolección de datos

Documentales

Fichaje

Según Arce (citado en Galvez Quilca, 2019) indica que “nos posibilita hacer un conveniente orden de la bibliografía, además este procedimiento consiste en sustraer, juntar, y ordenar de manera lógica los datos e ideas más importantes” (pág. 64). Bajo este procedimiento se concentra la data, recopilamos las ideas y después se organiza por medios de una ficha digital o física. Para resumir, lo cual se hace es el traspaso de los datos e información recolectada de libros, revistas u otro archivo digital o físico a un archivo o instrumento determinado como ficha.

Lectura Científica

Necesita de un lector(a) activo, que use sus conocimientos para lograr un raciocinio sobre lo leído. Un lector que logre diferenciar lo real o falso de los contenidos y, si es viable, producir otras ideas con base a las leídas en el documento. Este tipo de lectura es imprescindible para promover descubrimientos y novedades en las diferentes disciplinas, también es recomendable empezar con el resumen del archivo porque es la única manera de verificar que el documento contenga el tema deseado.

Los instrumentos

Encuesta

Es una técnica de recolección de datos por medio de la aplicación de un cuestionario a una muestra de personas. Mediante las encuestas se conoce las opiniones, las actitudes, y los comportamientos de los habitantes, consumidores, volantes, entre otros. La encuesta es el instrumento más usado y conocido una vez que se desea conseguir precisión y representatividad partiendo de manera directa de las consideraciones personales y no estructurales, para lograr una conclusión sobre los sujetos insertos en ellos (Arenal Laza, 2019).

La presente encuesta que se adjunta en el anexo 5 se aplicó de manera online a docentes de la Facultad de Ciencias Matemáticas y Físicas, para corroborar que tan aceptable es el desarrollo del presente proyecto.

Entrevista

Delgado (citado en Cedeño Pinargote & Chinga Quimis, 2020) indica: “Es una conversación entre dos o más sujetos que posees objetivos definidos, en donde se ejercen dos roles: el entrevistador y el entrevistado.

La presente entrevista que se adjunta en el anexo 5 se aplicó mediante diferentes reuniones por la plataforma zoom, a las profesionales logopedas, Lcda. Tairy Pérez y Lcda. Zuleika Quiñonez respectivamente, quienes con su criterio profesional respondieron satisfactoriamente la entrevista aplicada.

Técnicas estadísticas para el procesamiento de la información.

Técnicas para el Procesamiento y Análisis de Datos.

En este apartado se muestra el análisis de datos mediante las gráficas, tabulaciones, frecuencias absolutas y relativas, de la entrevista aplicada a las especialistas logopedas dentro de la ciudad de Guayaquil.

Conocimiento

Pregunta 1: ¿Qué test es el que más utiliza para diagnosticar estos problemas de fonemas?

Tabla 17

Pregunta 1: ¿Qué test es el que más utiliza para diagnosticar estos problemas de fonemas?

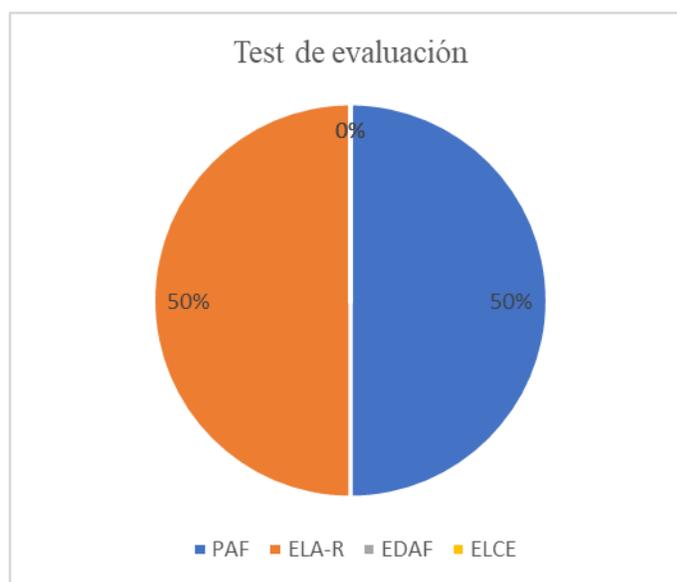
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	PAF	1	50,0	50,0	50,0
	ELA-R	1	50,0	50,0	100,0

EDAF	0	00,0	00,0	100,0
ELCE	0	00,0	00,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 1 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 7

Pregunta 1: Análisis gráfico de la pregunta número 1 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 50% indica que, usa la prueba de evaluación PAF, y el otro 50% indica usar el test ELA-R. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: Según los resultados obtenidos se afirma que estos especialistas utilizan las pruebas de evaluación según su criterio, en estos casos las más favorables para cada entrevistado fue el test ELA-R y el test PAF, pues en cada una de estas pruebas contienen los fonemas que se evalúa en cada paciente, con el cual servirá de ayuda para la construcción de la base de datos para el presente proyecto.

Factores que más influyen

Pregunta 2: Según su experiencia, ¿Cuáles son los factores que más influyen en estos trastornos de lenguaje?

Tabla 18

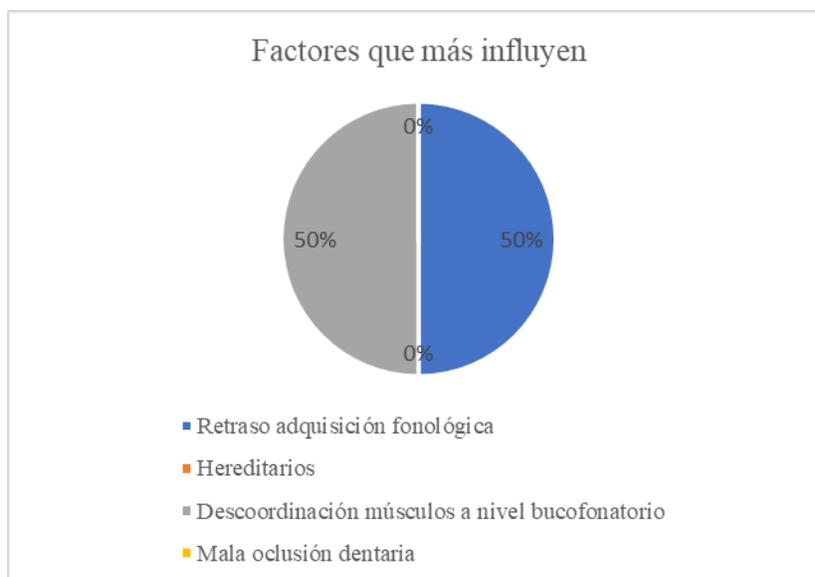
Pregunta 2: Según su experiencia, ¿Cuáles son los factores que más influyen en estos trastornos de lenguaje?

Factores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Retraso adquisición fonológica	1	50,0	50,0	50,0
Hereditarios	0	00,0	00,0	50,0
Descoordinación músculos a nivel bucofonatorio	1	50,0	50,0	100,0
Mala oclusión dentaria	0	00,0	00,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 2 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 8

Pregunta 2: Análisis gráfico de la pregunta número 2 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 50% indica que, como primer factor el retraso de adquisición fonológica, y el otro 50% indica la descoordinación de los músculos a nivel bucofonatorio. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: Los resultados que se logran visualizar en la gráfica muestran que la mayoría de los casos los factores que influyen en estos problemas de dislalia son por retraso de adquisición fonológica y por descoordinación de músculos a nivel bucofonatorio los cuales no permiten que los pacientes puedan articular los fonemas o palabras de manera correcta.

Frecuencia de pacientes

Pregunta 3: ¿Con que frecuencia se presentan pacientes con este tipo de trastornos?

Tabla 19

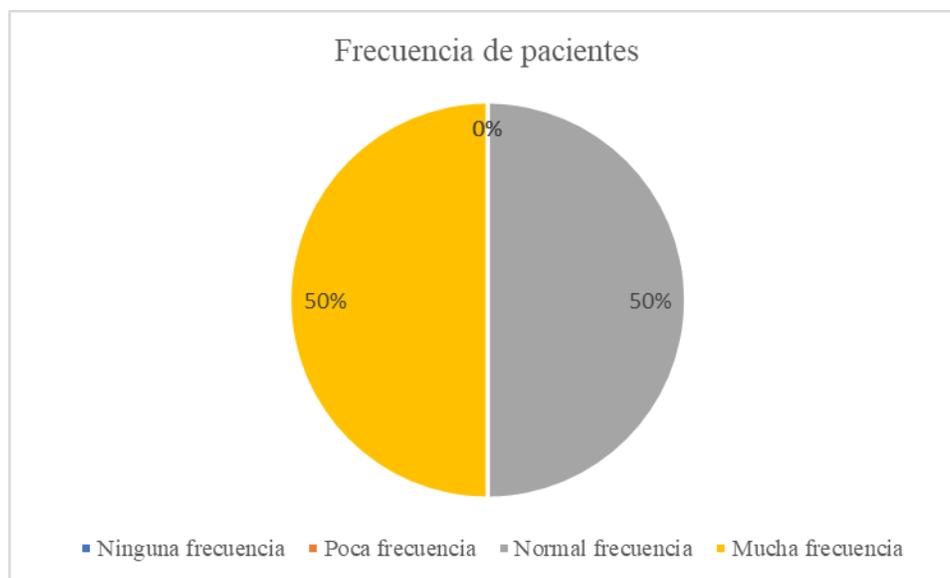
Pregunta 3: ¿Con que frecuencia se presentan pacientes con este tipo de trastornos?

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Ninguna frecuencia	0	00,0	00,0	00,0
Poca frecuencia	0	00,0	00,0	00,0
Válidos Normal frecuencia	1	50,0	50,0	50,0
Mucha frecuencia	1	50,0	50,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 3 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 9

Pregunta 3: Análisis gráfico de la pregunta número 3 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 50% indica que, es de normal frecuencia los pacientes con este tipo de trastornos, y el otro 50% indica que se presenta con mucha frecuencia pacientes con este tipo de trastornos. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: Se logra visualizar en el gráfico presentado que existe una igualdad de respuestas favorables a los enunciados denominados como “Mucha Frecuencia” y “Normal Frecuencia”, reflejando así un gran impacto en cuanto a la presencia de este tipo de trastornos dislálicos en los pacientes, dejando claro que este tipo de problemas deben ser tratados con mayor brevedad posible para poder disminuir el número de pacientes que lo padecen.

Consideración

Pregunta 4: ¿Cuáles son los fonemas más recurrentes que se presentan en sus pacientes?

Tabla 20

Pregunta 4 ¿Cuáles son los fonemas más recurrentes que se presentan en sus pacientes?

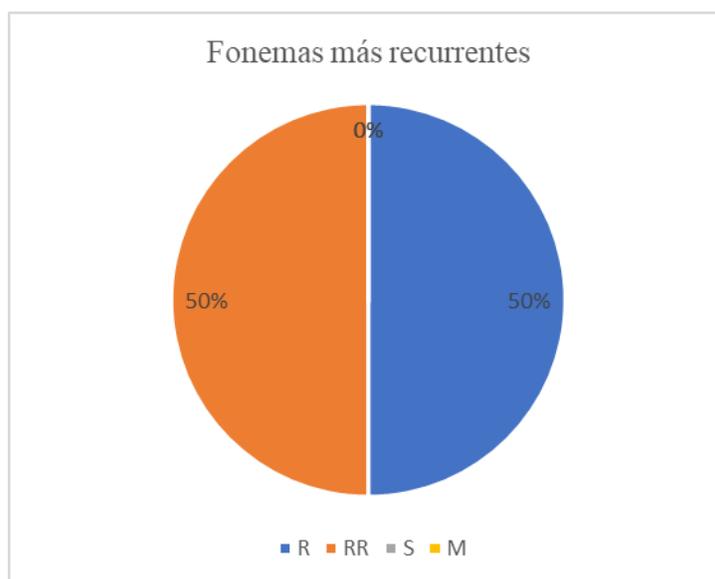
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	R	1	50,0	50,0	50,0
	RR	1	50,0	50,0	100,0

S	0	00,0	00,0	100,0
M	0	00,0	00,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 4 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 10

Pregunta 4: Análisis gráfico de la pregunta número 4 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 50% indica que, el fonema R es recurrente, y el otro 50% indica que el fonema RR. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: Cómo se puede visualizar en la gráfica según los especialistas logopedas experimentados en el tema, indican como resultado que los fonemas que padecen con mayor frecuencia los pacientes han sido el fonema “R” y el “RR”, lo cual va de la mano con de descoordinación de los músculos bucofonatorios, impidiéndoles así poder pronunciar las diversas palabras que conlleven este fonema, y como consecuencia a esto se desencadena el bullying hacia estos pacientes.

Perjudicial

Pregunta 5: ¿Qué tan perjudicial puede llegar a ser este tipo de trastornos en los pacientes?

Tabla 21

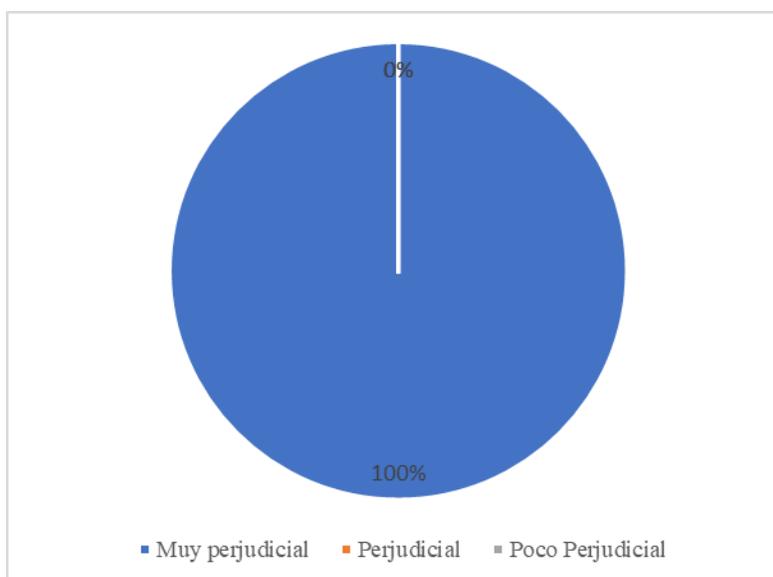
Pregunta 5: ¿Qué tan perjudicial puede llegar a ser este tipo de trastornos en los pacientes?

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos				
Poco perjudicial	0	00,0	00,0	00,0
Perjudicial	0	00,0	00,0	00,0
Muy perjudicial	2	100,0	100,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 5 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 11

Pregunta 5: Análisis gráfico de la pregunta número 5 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 100% indica que, es muy perjudicial este tipo de trastornos en los pacientes. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: Los resultados reflejados en el gráfico relacionado a la pregunta 5 según los entrevistados, nos indican que este tipo de trastornos puede llegar a ser “muy perjudicial” para la persona que lo padezca, pues impide al desarrollo educativo y social en la vida del paciente, así mismo como también se menciona es uno de las primeras causas por las cuales sufren discriminación o bullying por no poder pronunciar de manera correcta ciertos fonemas.

Asertividad

Pregunta 6: ¿Qué grado de asertividad brinda la realización de los test en los pacientes con dislalia?

Tabla 22

Pregunta 6: ¿Qué grado de asertividad brinda la realización de los test en los pacientes con dislalia?

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Nada de asertividad	0	00,0	00,0	00,0
Poca asertividad	0	00,0	00,0	00,0
Válidos Asertividad	0	00,0	00,0	00,0
Mucha asertividad	2	100,0	100,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 6 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 12

Pregunta 6: Análisis gráfico de la pregunta número 6 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 100% indica que, es de mucha asertividad la realización de los test en los pacientes con dislalia. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: El presente gráfico que está relacionado con la pregunta 6, según el criterio de los especialistas entrevistados indicaron que los test que ellos utilizan en los pacientes les brinda un alta asertividad, lo cual es confiable y recomendable para su uso, recalcando así mismo que los test más usados son el ELA-R y el PAF.

Proceso

Pregunta 7: ¿Cuál es el proceso a seguir para ayudar a mejorar este tipo de trastornos?

Tabla 23

Pregunta 7: ¿Cuál es el proceso a seguir para ayudar a mejorar este tipo de trastornos?

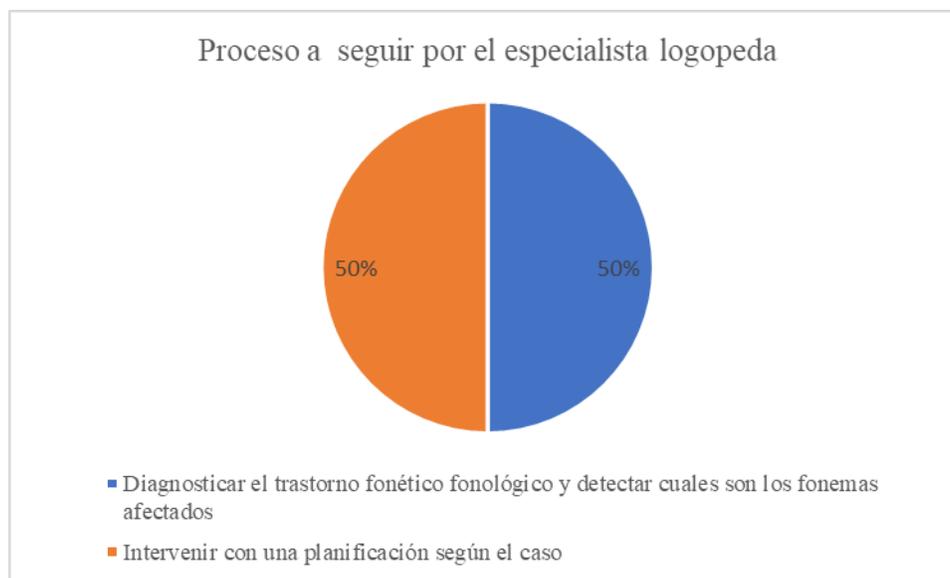
	Frecuenci a	Porcent aje	Porcentaje válido	Porcentaje acumulado
Diagnosticar el trastorno fonético fonológico y detectar cuales son los fonemas afectados	1	50,0	50,0	50,0

Intervenir con una planificación según el caso	1	50,0	50,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 7 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 13

Pregunta 7: Análisis gráfico de la pregunta número 7 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 50% indica que, diagnosticar el trastorno fonético fonológico y detectar cuales son los fonemas afectados es uno de los procesos y el otro 50% se complementa en intervenir con una planificación según el caso. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: Los resultados reflejados en el gráfico de la pregunta número 7 indican que es recomendable tener muy claro cuál es el fonema que mas afecta al paciente y a su vez también es necesario realizar una planificación con un tratamiento adecuado que ayude al paciente a poder erradicar estos problemas y así poder tener un estilo de vida normal.

Tiempo

Pregunta 8: ¿Cuál es el tiempo aproximado para poder tratar este tipo de trastornos?

Tabla 24

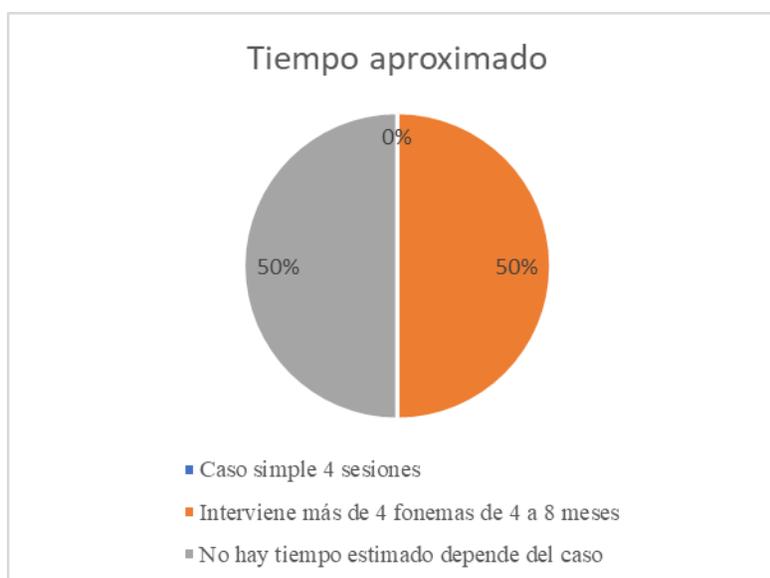
Pregunta 8: ¿Cuál es el tiempo aproximado para poder tratar este tipo de trastornos?

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Caso simple 4 sesiones	0	00,0	00,0	00,0
Interviene más de 4 fonemas de 4 a 8 meses	1	50,0	50,0	50,0
No hay tiempo estimado depende del caso	1	50,0	50,0	100,0
Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 8 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 14

Pregunta 8: Análisis gráfico de la pregunta número 8 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 50% indica que, si intervienen más de 4 fonemas un tiempo de 4 a 8 meses y el otro 50% indica que no hay tiempo estimado depende del caso. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: El resultado representado en el gráfico de la pregunta 8 refleja de forma muy precisa que el tiempo estimado es muy incierto debido a que cada paciente puede tardar mas tiempo o menos en referente a otro paciente y a su vez también influyen la cantidad de fonemas que se

presente estos inconvenientes, por esta razón según los expertos logopedas no se puede precisar con mayor exactitud un tiempo fijo para poder brindarle una mejoría en torno a estos problemas de dislalia.

Uso

Pregunta 9: ¿Si existiera un algoritmo para diagnóstico estaría dispuesto a usar esta herramienta, para que le genere un diagnóstico del trastorno dislálico?

Tabla 25

Pregunta 9: ¿Si existiera un algoritmo para diagnóstico estaría dispuesto a usar esta herramienta, para que le genere un diagnóstico del trastorno dislálico?

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
	Si	2	100,0	100,0	100,0
Válidos	No	0	00,0	00,0	100,0
	Total	2	100,0	100,0	

Nota: En esta tabla se muestran los valores absolutos y relativos correspondientes al proceso de tabulación de la Pregunta 9 aplicada a la entrevista a 2 individuos seleccionados para la investigación.

Figura 15

Pregunta 9: Análisis gráfico de la pregunta número 9 de la entrevista.



Nota: De un total de 2 especialistas entrevistados se observa que el 100% indica que, si usase el algoritmo para diagnóstico para dislalia funcional. La elaboración es propia y la fuente obtenida de la entrevista realizada de forma online.

Análisis: Los resultados en el gráfico de la pregunta número 9 reflejan los especialistas logopedas están en total acuerdo con un algoritmo para diagnóstico que les brinde la ayuda para obtener resultados automatizados y en el menor tiempo posible con la finalidad de que se pueda llevar un mejor control a los pacientes que se realicen este tipo de mejorías.

Metodología de desarrollo del prototipo

Metodología Kanban

El concepto Kanban ha pasado a conformar parte de las denominadas metodologías ágiles, cuyo objetivo es gestionar de forma general cómo se van completando las labores.

Kanban es un vocablo japonés que significa “tarjetas visuales”, donde Kan es “visual”, y Ban corresponde a “tarjeta”. (Gilbets, 2020).

Esta metodología se la utiliza para poder tener una mejor organización al momento de enlistar las actividades a realizar durante el proyecto y así poder terminar en el tiempo indicado cada actividad y evitar los cuellos de botellas.

Fases de la metodología

La metodología Kanban se divide en tres fases, las cuales son:

Request. - Planificación de las actividades a realizar al momento de desarrollar el algoritmo.

In Progress. - Las actividades que están en progreso durante las etapas del proyecto.

Done. - Las actividades que ya se han realizado durante la realización del proyecto.

Fase 1

❖ Diseño Base de Datos - Dataset

Se pudo obtener acceso a una base de datos dentro del proyecto FCI de la Facultad de Ciencias Matemáticas y Físicas, levantada con 50 datos reales de pacientes con diferentes tipos de dislalias y diferentes atributos, para la construcción del algoritmo para diagnóstico este proyecto se ha enfocado en la dislalia funcional por lo cual para la selección de los atributos/variables más significativos de la base de datos realizamos lo siguiente:

1. Investigaciones dentro de artículos científicos, revistas científicas, libros sobre dislalia, fichas logopédicas que comprenden las pruebas de evaluación ELA-R, PAF, entrevista a un especialista logopeda el cual con su criterio nos ayudó a descartar y seleccionar las variables más óptimas en la toma de decisión del diagnóstico de la dislalia funcional.

2. Lo siguiente fue crear un dataset en Excel a partir de la base de datos, el cual está conformado por:

Variables: [M -P- B- D- T- Z- F- S- N- L- R- RR- LL- Ñ- K- G- J- X- BL- CL- FL- GL- PL- BR- CR- DR- PR- GR]

Datos para cada variable del déficit de fonemas dentro de dislalia funcional:

Sustitución inicio, medio, final [1.1 – 1.2 – 1.3]

Omisión inicio, medio, final [2.1 – 2.2 - 2.3]

Distorsión inicio, medio, final [3.1 – 3.2 – 3.3]

Inversión inicio, medio, final [4.1 – 4.2 – 4.3]

Adición inicio, medio, final [5.1 - 5.2 – 5.3]

y los datos entorno a las pruebas realizados a los pacientes que ayudarían al proceso del diagnóstico de la dislalia funcional.

3. Por último, contando con los datos y variables óptimas en este dataset realizado en Excel, se almaceno un archivo con extensión .csv para su posterior uso en el algoritmo construido en Python.

Fase 2

❖ Simulación de Datos

Una vez selecciona las variables y creado el nuevo dataset, se resolvió que los datos en la base real de pacientes con dislalia funcional eran muy pocos, para lo cual se realizó una simulación de 1002 datos, tomando los datos reales y por consiguiente aumentando más casos para el dataset.

Las herramientas utilizadas para la simulación de estos datos fueron dos: Microsoft Excel el cual tiene integrado Microsoft Visual Basic para aplicaciones, herramienta mediante la cual se generarán los 1002 datos de forma aleatoria. Para llevar a cabo esta simulación se utilizará el método Monte Carlo ayudados de la herramienta @Risk 8.1 que es un complemento de Excel usado para la aplicación de simulación.

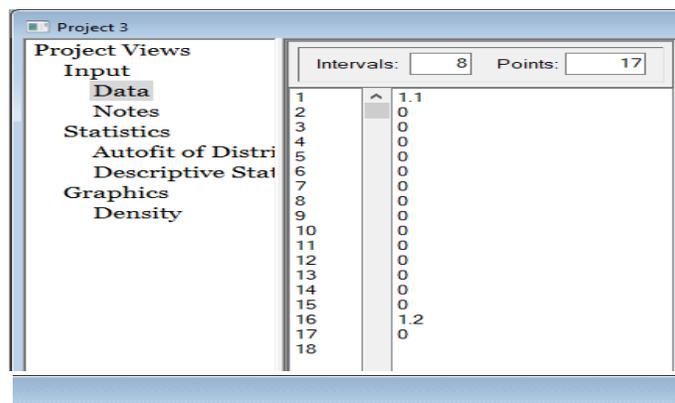
❖ Simulación Monte Carlo

El propósito primordial de la simulación de Montecarlo es intentar emular la conducta de variables reales para, en el tamaño de lo viable, examinar o pronosticar cómo van a cambiar. La simulación Monte Carlo es un instrumento estadístico que posibilita la modelación de resultados conforme con la conducta histórico de los datos y su posibilidad de ocurrencia, (Salazar & Alzate, Aplicación de la simulación Monte Carlo en la proyección del estado de resultados. Un estudio de caso, 2018)

Cabe aclarar que para poder utilizar el método Monte Carlo se debe conocer las distribuciones de cada una de las variables del dataset, esto lo logramos haciendo uso de la herramienta STAT::FIT, esta herramienta tecnológica nos permite la entrada de mínimo 10 y máximo 50 datos, dado el caso que el dataset contiene los datos reales de pacientes estos son ingresados en la herramienta por variable como indica la figura 16.

Figura 16

STAT::FIT. Distribución de la variable M



descriptive statistics	
data points	17
minimum	0
maximum	1.2
mean	0.135294
median	0
mode	0
standard deviation	0.38233
variance	0.146176
coefficient of variation	282.592
skewness	2.62054
kurtosis	4.33162

Nota: En esta grafica se muestra información sobre la variable M del dataset, donde se puede observar, datos, estadística descriptiva. La elaboración es propia y las fuentes obtenidas a través de las investigaciones realizadas.

Figura 17

Simulación de datos. Microsoft Visual Basic.

```

Dim Beneficio As Double

c = Val(InputBox("Iteracciones", , 1000))

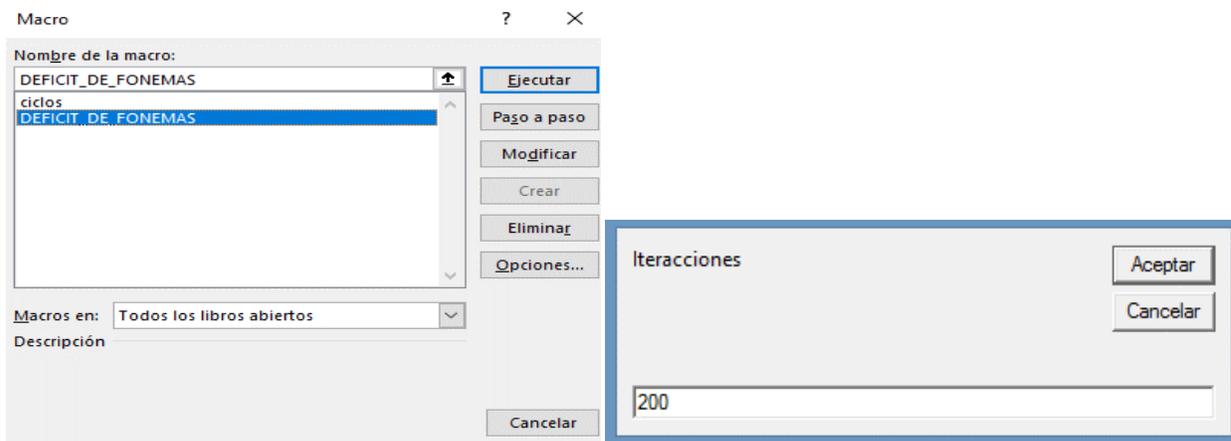
For i = 2 To c

    M = Range("C3")
    P = Range("C4")
    B = Range("C5")
    D = Range("C6")
    T = Range("C7")
    Z = Range("C8")
    F = Range("C9")
    S = Range("C10")
    N = Range("C11")
    L = Range("C12")
    R = Range("C13")
    RR = Range("C14")
    LL = Range("C15")
    Ñ = Range("C16")
    K = Range("C17")
    G = Range("C18")
    J = Range("C19")
    X = Range("C20")
    BL = Range("C21")
    CL = Range("C22")
    FL = Range("C23")
    GL = Range("C24")
    PL = Range("C25")
    BR = Range("C26")
    CR = Range("C27")
  
```

Nota: En este grafico se muestra el código dentro del editor Microsoft Visual Basic, ejecutando el proceso para la obtención de los 2000 datos. La elaboración es propia y las fuentes obtenidas a través de las investigaciones realizadas.

Figura 18

Iteraciones por cada una de la clasificación de déficit de fonemas.



Nota: Este grafico muestra la cantidad de iteraciones necesarias (200) por cada una de las clasificaciones del déficit de los fonemas hasta obtener los 1002 datos. La elaboración es propia y la fuente es obtenida a través de las investigaciones.

Figura 19

Simulación Montecarlo. Resultado final de la simulación.

		Edad	M	P	B	D	T	Z	F	S	N	L	R	RR	LL	Ñ	K	G	J	X	BL	CL	FL	GL	PL	BR	CR	DR	PR	GR
edad		6	0	0	0	0	0	1.1	0	2.2	0	0	3.1	3.1	4.2	0	0	0	0	1.1	0	0	0	0	0	3.1	3.1	3.1	3.1	3.1
M	0,78	5	0	2.1	2.1	0	0	0	0	1.1	2.2	1.1	2.2	1.1	5.1	2.3	2.3	0	2.3	2.3	1.1	4.2	2.1	2.1	2.1	1.1	2.1	2.1	1.1	4.2
P	0,00	7	0	2.1	1.1	2.1	2.1	2.1	0	2.2	5.1	2.2	1.1	2.2	0	1.1	2.3	1.1	2.3	1.1	1.1	5.1	2.1	1.1	2.1	1.1	2.1	2.1	1.1	2.1
B	0,00	8	0	2.1	2.1	1.1	1.1	2.1	2.3	1.1	2.2	4.2	2.2	1.1	0	1.1	2.3	2.3	2.3	1.1	2.3	1.1	2.1	2.1	1.1	2.1	2.1	2.1	1.1	2.1
D	0,00	9	0	0	0	0	0	0	0	1.2	0	0	1.2	1.2	1.2	5.1	2.1	2.1	2.1	0	1.2	1.2	0	0	2.1	1.2	1.2	1.2	1.2	1.2
T	0,00	5	2.2	2.1	0	2.1	2.1	2.1	1.2	0	0	0	0	1.2	1.3	0	2.1	2.1	0	1.2	0	5.1	2.1	2.1	2.1	1.2	1.2	1.3	1.3	1.3
Z	1,00	6	0	0	0	0	0	0	0	0	0	0	3.1	3.1	1.3	2.3	1.1	4.2	1.2	1.3	0	0	2.1	2.1	0	3.1	3.1	3.1	3.1	3.1
F	1,00	7	1.3	2.1	2.1	1.1	5.1	2.1	1.3	2.2	2.2	4.2	2.2	1.3	1.3	0	1.1	1.1	0	1.3	1.3	2.3	1.1	4.2	1.2	1.3	2.1	2.1	1.3	2.1
S	0,00	8	2.2	0	1.2	1.2	2.1	3.1	0	2.2	1.2	1.2	1.2	3.1	0	1.2	0	0	0	1.2	0	0	1.1	1.1	0	1.2	2.1	2.1	2.1	0
N	0,00	9	0	0	0	1.3	0	2.1	0	0	1.3	0	0	2.2	0	0	1.3	0	1.1	2.3	1.1	2.3	1.1	4.2	1.2	2.1	2.1	1.2	1.3	1.3
L	0,00	5	2.2	2.1	0	1.3	0	2.1	0	0	0	0	0	2.2	2.2	1.3	2.3	0	1.1	0	1.1	0	1.1	1.1	0	1.1	1.2	0	0	4.2
R	0,00	6	2.2	0	0	1.1	0	1.1	0	1.1	0	2.2	0	2.2	0	2.3	0	0	0	1.1	0	1.1	1.2	0	1.2	1.2	2.1	1.2	0	1.2
RR	1,00	5	0	1.3	1.3	0	0	2.1	0	0	1.3	0	1.2	0	1.3	2.1	0	0	0	1.2	0	0	1.2	0	1.1	2.1	3.1	1.1	2.1	1.2
LL	1,00	6	0	1.1	1.1	2.1	0	0	1.1	1.1	1.1	1.3	1.3	1.3	2.2	2.1	1.2	1.2	2.3	2.3	0	0	0	1.2	1.2	0	0	0	2.1	1.2
Ñ	0,00	5	0	1.2	3.1	0	0	0	1.3	0	1.2	2.2	1.2	2.2	0	4.2	0	0	1.2	1.2	0	1.1	1.1	0	0	2.1	0	0	1.2	1.3
K	0,00	5	0	0	0	0	0	1.3	2.1	0	0	0	1.3	1.3	2.2	1.1	0	0	0	1.3	0	0	0	0	0	1.3	1.3	1.1	1.1	1.1
G	0,00	5	0	2.1	1.1	2.1	4.1	2.1	2.1	2.2	1.1	2.2	1.1	2.2	2.2	0	2.3	1.1	2.3	2.3	2.3	2.3	2.1	2.1	2.1	1.1	2.1	2.1	1.1	2.1
J	0,00	6	2.2	1.1	1.1	2.1	2.1	1.1	1.1	5.1	2.2	1.1	1.1	0	0	2.3	1.3	0	1.3	1.3	5.1	5.1	2.1	2.1	2.1	1.3	1.3	1.3	1.2	1.2
X	1,00	5	0	2.1	0	2.1	2.1	2.1	0	0	0	0	0	1.2	5.1	2.1	2.1	2.1	2.3	1.2	5.1	2.1	2.1	2.1	0	1.1	1.1	1.1	1.1	1.1
BL	0,00	7	1.3	2.1	0	1.1	0	1.1	0	2.2	0	0	2.2	1.1	0	2.1	2.1	0	2.3	1.1	0	2.1	2.1	0	0	1.1	1.1	1.1	1.1	1.2
CL	0,00	8	2.2	0	0	0	0	0	0	0	0	0	3.1	3.1	2.3	1.1	4.2	1.2	0	1.2	2.3	1.1	4.2	1.2	0	3.1	3.1	3.1	3.1	
FL	0,00	9	0	2.1	0	1.3	0	1.3	0	2.2	0	0	2.2	1.3	0	1.1	1.1	0	2.3	1	0	1.1	1.1	0	0	1.1	1.1	1.2	1.2	

Nota: Se presenta los 1002 datos simulados correspondientes a las 5 clasificaciones de déficits de fonemas dentro de la dislalia funcional. La elaboración es propia y las fuentes obtenidas de las investigaciones realizadas.

Elección del algoritmo.

Según el análisis presentado en el cap. II, donde se midieron a través de indicadores varios algoritmos de machine learning y el algoritmo basado en reglas, se llegó a la conclusión que al no obtener un resultado positivo en la precisión mediante algoritmos de machine learning se eligió algoritmos basados en reglas el cual va a ser utilizado en el desarrollo de este prototipo ya que con ello se pudo satisfacer el objetivo principal que es generar un diagnóstico a partir del dataset el cual representa las fichas logopédicas.

Fase 3

❖ Diseño Algoritmo

- **Base de datos:** se trabajó con 1002 datos alojados en el dataset (.csv)
- **Preparación entorno de trabajo:** Se instalaron herramientas como el editor de código Visual Studio Code que nos permitió con su enorme cantidad de extensiones gratuitas poder desarrollar en el lenguaje Python, de forma amigable el algoritmo para diagnóstico.
- **Creación entorno de trabajo:** para que en el entorno de trabajo se pueda construir el algoritmo se instalaron las extensiones de las librerías necesarias.

❖ Inventariar el algoritmo

En esta sección se detalla todo lo que contiene el prototipo de algoritmo basado en reglas inventariado mediante tablas y su respectiva descripción.

Librerías

Tabla 26

Librerías dentro del algoritmo

Librerías	Descripción
Pandas	Destinada al análisis de los datos
Time	Retorna el tiempo expresado como número de coma flotante
Tkinter	Interfaz gráfica para Python

Nota: La grafica indica las librerías dentro del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Variables

Primer bloque de código: Aquí se indican las variables de entrada, almacenamiento de los datos y cabecera de datos tomada del dataset de entrada, las variables indicadas en la tabla 25 solo toman valores de entrada y almacenamiento dentro del código.

Tabla 27

Variables dentro del algoritmo

Variables	Descripción
Data	Almacena el dataframe con los datos y etiquetas del archivo .csv
datos	Almacena una representación numpy de la trama de datos
Columnas	Almacena las etiquetas de los datos

Nota: La grafica indica las variables de entrada y almacenamiento dentro del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Segundo bloque de código: Aquí se encuentran las variables del módulo tk, el cual nos va a proporcionar la información visual el posicionamiento y control de la pantalla, cabe destacar que estas variables no toman otros valores, su objetivo es la creación de una interfaz amigable para la interacción usuario-algoritmo.

Tabla 28*Variables para información visualización*

Variables	Descripción
Self	Objeto de clase Example, recibe un objeto cuadro tkinter y sirve para declarar todos los elementos de la interfaz grafica
Prompt	Muestra en pantalla la espera de las indicaciones de órdenes. Es de tipo etiqueta de tkinter.
entry	Usada para aceptar la cadena de texto Tipo: Entrada de tkinter
Output	Muestra la salida del algoritmo. Se inicia con una cadena vacía. Tipo etiqueta de Tkinter.
Submit	Al dar clic envía la entrada del usuario (entry) para que sea procesada por el algoritmo en la función diagnosticar. Tipo botón de Tkinter.
Finalizar	Crea un botón finalizar que al ser presionado termina el proceso con el comando root.quit
Title	Le da el nombre a la pantalla.

Nota: La grafica indica las variables de entrada y almacenamiento dentro del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Tercer bloque de código: dentro de este fragmento de código se pueden especificar cada una de las variables que se ubican dentro de la función diagnosticar. La cuales cumplen varios roles como se especifica en la siguiente tabla.

Tabla 29*Variables dentro de la función diagnosticar*

Variables	Descripción
Paciente	Almacena el numero ingresado por el usuario
datos	Almacena una representación numpy de datos en el archivo .csv
inicio	Almacena el tiempo en que se inicia la función diagnosticar
x	Recorre la variable datos y devuelve una lista por cada fila(paciente) del .csv

A	Almacena una sublista de la variable x empezando en la segunda posición
Y	Recorre la sublista a para obtener cada valor por fila
Sustitución	Almacena los datos (fonemas) que cumplen con los condicionales para sustitución que se comparen durante el recorrido de los datos del dataset e indicará la posición afectada.
Omisión	Almacena los datos (fonemas) que cumplen con los condicionales para omisión que se comparen durante el recorrido de los datos del dataset e indicará la posición afectada.
Distorsión	Almacena los datos (fonemas) que cumplen con los condicionales para distorsión que se comparen durante el recorrido de los datos del dataset e indicará la posición afectada.
Inversión	Almacena los datos (fonemas) que cumplen con los condicionales para inversión que se comparen durante el recorrido de los datos del dataset e indicará la posición afectada.
Adición	Almacena los datos (fonemas) que cumplen con los condicionales para adición que se comparen durante el recorrido de los datos del dataset e indicará la posición afectada.
Total	Almacena el conjunto de datos guardados en los arrays sustitución, omisión distorsión inversión y adición
newlist	Almacena los valores de los arrays eliminando valores vacíos.
Final	Almacena el tiempo en que termina la función diagnosticar
Tiempo	Almacena el tiempo que demora la función diagnosticar
Result	Almacena un string de los valores en newlist listos para su presentación en pantalla
Submit	Tipo botón tkinter al ser presionado llama a la función limpiar que elimina los valores presentados e ingresados para realizar un nuevo diagnostico

Nota: La tabla indica las variables dentro de la función diagnosticar del algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Tipos de variables

Tabla 30

Tipos de variables

Tipos	Variables
Arrays	data – datos – columnas – a – x – total – newlist – Sustitucion - Omision – distorsión. inversión – Adicion. Sirven para almacenar el diagnóstico del cliente.
Int	Paciente, almacena el número del paciente a diagnosticar
float	Variable y, almacena la representación del diagnóstico del paciente de cada fonema.
String	Result, almacena el texto presentado en pantalla con el diagnóstico del paciente.
File	Objeto archivo f, sirve para crear y escribe un archivo txt con el diagnostico.

Nota: La grafica indica los tipos de variables. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Funciones dentro del prototipo de algoritmo

Tabla 31

Funciones dentro del algoritmo para diagnostico

Funciones	Descripción
init	Inicializa la pantalla y sus elementos
Diagnosticar	Se cuenta con una función llamada diagnosticar ()
Len	Retorna el número de ítems en un contenedor
Time	Retorna el tiempo actual en segundos
Open	Escribe el archivo de archivo y establece varias banderas de acuerdo con las banderas y posiblemente su modo según el modo. El modo predeterminado es 0777 (octal).
limpiar	Recibe las variables de los elementos de la pantalla y los inicializa a su valor original.
Round	Redondea un número a la cantidad de decimales específicos

Nota: La grafica indica las siete funciones dentro del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas

Métodos dentro del prototipo de algoritmo

Tabla 32

Métodos dentro del algoritmo

Métodos	Descripción
read_csv	Inicio del algoritmo, recibe la ruta del archivo .csv, el caracter separador (el predeterminado es ',') y la codificación del alfabeto latino, devuelve un nuevo dataframe con los datos y etiquetas del archivo .csv
Values	Devuelve los datos del dataframe datos ordenados en un arreglo numpy
Columns	Devuelve las etiquetas del archivo .csv ordenados en un array
Delete	Elimina el elemento del array columna en la posición especificada
Pack	Posiciona las variables tipo tkinter en el cuadro, recibe las posiciones: top, botton, left, right.
Tk_setPalette	Cambia el color de fondo de la pantalla
Configure	Actualiza el texto o el comando del elemento tkinter que lo invoca
Entry.Delete	Elimina el valor que ingresa el usuario
Entry.get	Devuelve el valor que el usuario ingresa para almacenarlo en la variable paciente
startswith	Devuelve verdadero si empieza con el prefijo especificado, falso en caso contrario.
Append	Anexar objeto al final de la lista
Split	Devuelve una lista de la palabra de la cadena, utilizando el carácter SEP como cadena delimitadora
Write	Escribe una cadena str en el archivo. No hay valor devuelto.
Mainloop	Corre el evento Tkinter loop que representa la pantalla

Nota: La grafica indica los métodos dentro del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Condiciones

Estas condiciones presentes dentro del algoritmo cumplen la función de reglas a cumplirse debidamente, para de esta manera realizar el diagnóstico respectivo a cada paciente en función de su ficha logopédica.

Tabla 33

Condicionales

Condiciones	Descripción
	Cuenta con siete condicionales. Validan que:
	El número ingresado sea un número positivo dentro del rango de diagnóstico de pacientes disponibles en el dataset.
	Que el número ingresado corresponde a la fila del diagnóstico del paciente en el dataset
If	Que el valor en el dataset represente el trastorno dislábico: sustitución, omisión, distorsión, inversión o adición
	Si existen arrays vacíos en la lista total.
	Que el valor del dataset represente la posición del trastorno: inicio, medio y final
	Si el módulo ejecutado es el programa principal
Else	Opción alterna si los valores no cumplen con las condicionales if

Nota: La grafica indica las condicionales dentro del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas

Bucles

Son aquellas instrucciones que se ejecutaran dentro del código hasta que la condición definida deje de cumplirse.

Tabla 34

Bucles dentro del código

Bucles	Descripción
	El algoritmo cuanta con ocho bucles.
	Recorre la lista datos para acceder a los diagnósticos por paciente
for	Recorre la lista a para obtener el valor de cada diagnóstico de un paciente
	Recorre los arrays de resultado sustitución, omisión, distorsión, inversión, adición para remplazar los valores 1,2,3 por las posiciones inicio, medio y final (5)
	Recorre la lista newlist para remplazar valores vacíos por la etiqueta 'ninguno'
If - else	Diecisiete bucles

Nota: La grafica indica los bucles dentro del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Procesos

Tabla 35

Procesos dentro del algoritmo

Procesos	Descripción
Primero	Entrada y lectura de los datos a partir del dataset
Segundo	Procesamiento de los datos mediante condiciones, bucles, reglas y función.
Tercero	Resultado de diagnóstico y escritura de archivo .txt

Nota: La grafica indica los bucles de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas

Entradas y salidas del algoritmo

Tabla 36

Entradas y salidas del algoritmo

		Descripción
Entrada	Dataset.csv	Construido con 28 variables de entrada (fonemas) y 1002 pruebas de evaluación (datos) denominadas fichas logopédicas.
Salida	Diagnostico	Arroja el diagnostico en base a las entradas cargadas al algoritmo, indicando los resultados de cada paciente a partir de la ficha logopédica.

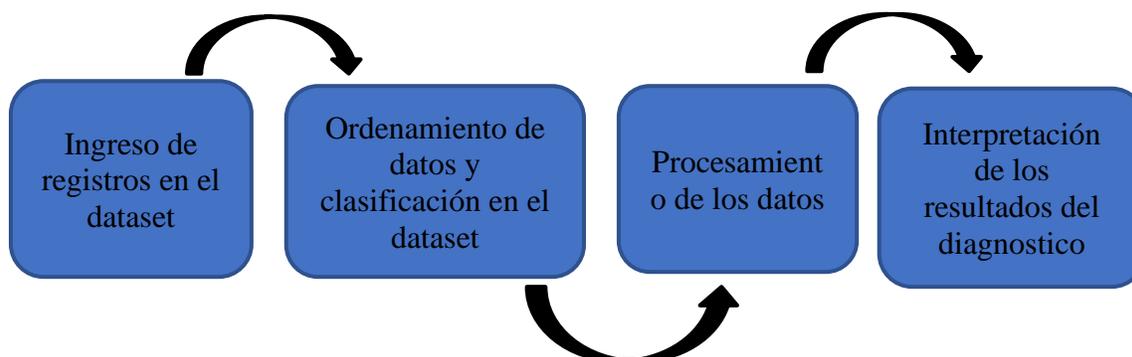
Nota: La grafica indica entrada y salida del prototipo de algoritmo basado en reglas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas

❖ Implementación

En los alcances del proyecto la implementación del algoritmo consiste en la construcción del algoritmo en Python, y el prototipo donde se introduzcan las entradas de los resultados más destacados de las fichas logopédicas y nos arroje un diagnóstico.

Figura 20

Proceso de implementación del algoritmo



Nota: Diagrama de flujo de implementación del algoritmo. La elaboración es propia y los datos de la investigación.

Figura 21

Implementación de algoritmo.

```

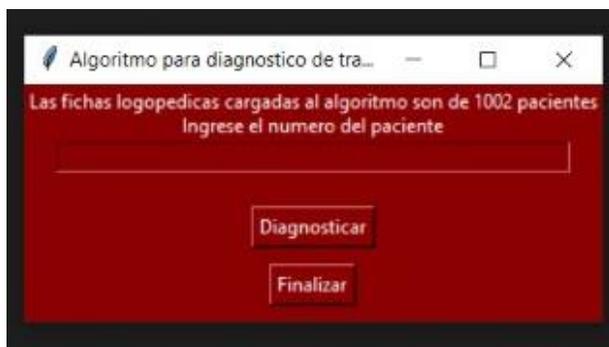
1 import tkinter as tk
2 import pandas as pd
3 import time
4 data = pd.read_csv('C:\\Users\\hp240\\Desktop\\juk\\PRUEBA.csv', sep=';', encoding='latin-1')
5 datos = data.values
6 columna = data.columns
7 columna=columna.delete(0)
8 columna=columna.delete(0)
9 columna = columna.values
10
11
12
13 class Example(tk.Frame):
14     def __init__(self, parent):
15         tk.Frame.__init__(self, parent)
16         self.prompt = tk.Label(self, text="Las fichas logopedicas cargadas al algoritmo son de " + str(len(datos)) + " pac
17         self.entry = tk.Entry(self)
18         self.output = tk.Label(self, text="")
19         self.submit = tk.Button(self, text="Diagnosticar", command = self.diagnosticar)
20         self.prompt.pack(side="top", fill="x")
21         self.entry.pack(side="top", fill="x", padx=20)
22         self.submit.pack(side="bottom")
23
24
25         self.output.pack(side="top", fill="both", expand=True)
26         tk.Button(root, text='Finalizar', command=root.quit).pack(side='bottom', padx=10, pady=10)
27         self.tk_setPalette('darkred')
28         root.title("Algoritmo para diagnostico de trastorno dislálico")
29
30
31
32     def limpiar(self):
33         self.submit.config(text="")

```

Nota: Editor de código Visual Studio Code, algoritmo construido en lenguaje Python. La elaboración es propia y los datos de la investigación.

Figura 22

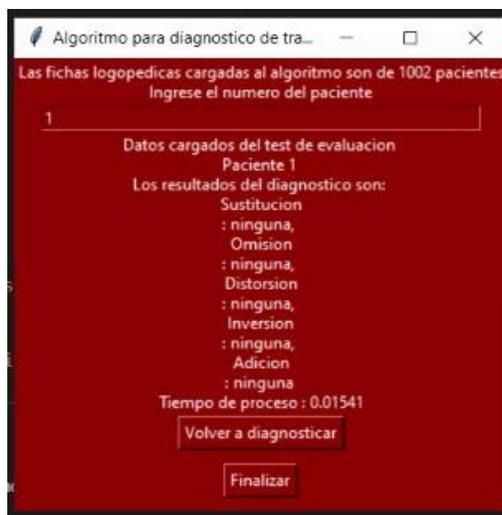
Pantalla de visualización al momento de ejecutar el algoritmo.



Nota: La siguiente figura nos muestra la pantalla de presentación la cual solicita que se digite el número del paciente al cual se le realizar el diagnóstico. La elaboración es propia y los datos de la investigación.

Figura 23

Pantalla de visualización del diagnóstico al paciente uno



Nota: La siguiente figura nos muestra la pantalla con el resultado del diagnóstico al paciente uno, también contiene dos botones para volver a realizar otro diagnóstico o finalizar la ejecución. La elaboración es propia y los datos de la investigación.

Tabla 37

Operacionalización de variables

OPERACIONALIZACIÓN DE VARIABLES

Comportamiento	Descripción
Funcionalidad	Permite que al momento de diagnosticar los resultados del diagnóstico se registren en un archivo txt.
Fiabilidad	Que el algoritmo sea de calidad y cómodo.
Usabilidad	Permite que se pueda manipular, controlar y entender el diagnóstico.
Eficiencia	Rendimiento del algoritmo para funcionar de forma correcta y eficiente.
Mantenibilidad	Manera del software que tenga para actualizarse y mejor el rendimiento de dicho algoritmo.
Portabilidad	Capacidad del algoritmo para diagnóstico logre adaptarse a diversos entornos y comunicar con los sistemas.

Nota: La elaboración es propia de los autores y la fuente es proporcionada por las investigaciones realizadas.

Pruebas Inducidas

En este apartado se medirá la efectividad del resultado del diagnóstico que nos otorgue el algoritmo, mediante pruebas inducidas en la cual se cambiarán los valores ingresados en el dataset y corroborar que el resultado sea el que ya tendremos conocimiento y así poder verificar que todas las reglas que se encuentran inmersas están correctamente formuladas dentro de este prototipo.

Prueba N°. 1

Figura 24

Dataset con datos de pacientes a diagnosticar

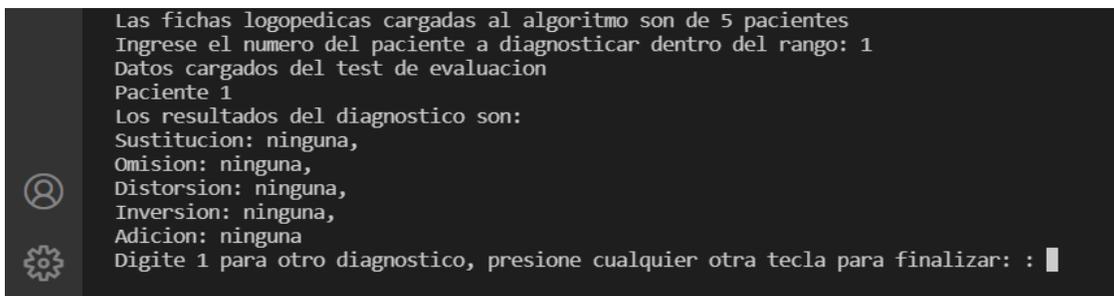
	A	B	C	D	E	F	G	H
1	Paciente	edad	M	P	B	D	T	Z
2	1	12	0	0	0	0	0	0
3	2	6	0	1.2	1.2	2.3	4.3	2.1
4	3	8	2.1	4.1	3.2	0	4.1	5.2
5	4	9	5.2	3.3	0	3.2	1.2	1.1
6	5	7	3.3	5.2	2.2	2.2	0	0
7								

Nota: En la gráfica se visualiza la matriz de datos a partir de la ficha logopédica. La elaboración es propia y los datos de la investigación realizada.

Primero se probará el algoritmo ingresando un Dataset con 5 pacientes como indica la figura 24, y los fonemas a evaluar serán solo 6, y se pedirá que nos arroje el diagnóstico del paciente 1, en el cual como resultado nos arrojará que no presenta ningún inconveniente en ninguna de los fonemas existentes como indica la figura 25.

Figura 25

Salida por consola del diagnóstico paciente 1



```

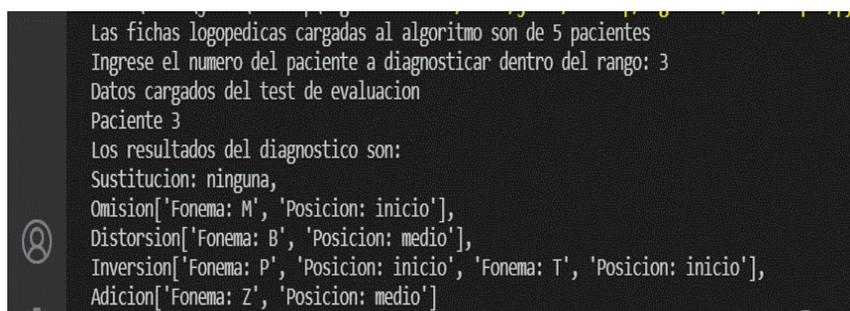
Las fichas logopedicas cargadas al algoritmo son de 5 pacientes
Ingrese el numero del paciente a diagnosticar dentro del rango: 1
Datos cargados del test de evaluacion
Paciente 1
Los resultados del diagnostico son:
Sustitucion: ninguna,
Omision: ninguna,
Distorsion: ninguna,
Inversion: ninguna,
Adicion: ninguna
Digite 1 para otro diagnostico, presione cualquier otra tecla para finalizar: : █
  
```

Nota: La grafica nos indica el diagnóstico del paciente uno. La elaboración es propia y los datos de la investigación realizada.

Como ya se había mencionado, el diagnóstico en el paciente 1 es el acertado, no presenta ningún tipo de trastornos en ninguna de los fonemas, se lo puede corroborar en la imagen, que indica cuales son la clasificación de los trastornos, pero en ninguno de ellos el paciente evaluado presenta inconvenientes. Así mismo se va a diagnosticar con el mismo dataset al paciente número 3 en el cual ya sabremos que deben presentar varios tipos de trastornos en el diagnóstico final.

Figura 26

Salida por consola del diagnóstico paciente 3



```

Las fichas logopedicas cargadas al algoritmo son de 5 pacientes
Ingrese el numero del paciente a diagnosticar dentro del rango: 3
Datos cargados del test de evaluacion
Paciente 3
Los resultados del diagnostico son:
Sustitucion: ninguna,
Omision['Fonema: M', 'Posicion: inicio'],
Distorsion['Fonema: B', 'Posicion: medio'],
Inversion['Fonema: P', 'Posicion: inicio', 'Fonema: T', 'Posicion: inicio'],
Adicion['Fonema: Z', 'Posicion: medio']
  
```

Nota: La grafica nos indica el diagnóstico del paciente tres. La elaboración es propia y los datos de la investigación realizada.

De la misma forma se puede visualizar como indica la figura 26 los resultados obtenidos son los esperados en cada uno de los fonemas a evaluar, es decir nos diagnostica que tipo de déficit de fonema presenta y en qué posición se encuentra la dificultad, en el cual se podría acotar que el algoritmo cumple con las expectativas propuestas.

Prueba N°. 2

Figura 27

Dataset con datos de pacientes a diagnosticar

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Paciente	edad	M	P	B	D	T	Z	F	S	N	L	R	RR
2	1	12	6.1	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3
3	2	6	0	0	0	0	0	0	0	0	0	0	0	0
4	3	8	2.1	4.1	3.2	0	4.1	5.2	0	1.1	2.1	0	0	3.1
5	4	9	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1
6	5	7	3.3	5.2	2.2	2.2	0	0	1.3	3.3	2.2	4.1	0	5.3
7														

Nota: En la gráfica se visualiza la matriz de datos a partir de la ficha logopédica. La elaboración es propia y los datos de la investigación realizada.

En esta segunda prueba se le ingresara al algoritmo un mismo dataset con 5 pacientes, y los fonemas (variables de entrada) a evaluar ya se incrementaron a 12, y en este caso se pedirá que nos arroje el diagnóstico del paciente 2, en el que obtendremos como resultado que no presenta ningún inconveniente en ninguna de los fonemas evaluados.

Figura 28

Salida por consola del diagnóstico paciente 2

```

Las fichas logopedicas cargadas al algoritmo son de 5 pacientes
Ingrese el numero del paciente a diagnosticar dentro del rango: 2
Datos cargados del test de evaluacion
Paciente 2
Los resultados del diagnostico son:
Sustitucion: ninguna,
Omision: ninguna,
Distorsion: ninguna,
Inversion: ninguna,
Adicion: ninguna

```

Nota: La grafica nos indica el diagnóstico del paciente dos con 12 variables de entrada. La elaboración es propia y los datos de la investigación realizada.

- **A Priori Analysis** - Este es un estudio teórico de un algoritmo. La eficiencia de un algoritmo se mide aceptando que todos los otros componentes, por ejemplo, la rapidez del procesador, son constantes y no poseen ningún impacto en la utilización.

- **A Posterior Analysis** - Este es un estudio experimental de un algoritmo. El algoritmo seleccionado se implementa por medio de lenguaje de programación. Esto se ejecuta en la PC de destino. En este estudio, se recopilan estadísticas reales como el tiempo de ejecución y el espacio requerido.

- **Time Factor** – El tiempo se mide contando el número de operaciones clave, como comparaciones en el algoritmo.

- **Space Factor** - El espacio se mide contando el espacio de memoria máximo solicitado por el algoritmo.

Tabla 38

Factor tiempo y espacio del algoritmo basado en reglas en máquina de investigadores

PC	Factor Tiempo		Factor Espacio
	Mínimo	Máximo	
Maquina 1: Philco Jorge	0,03125	0,0625	A razón de 4 Gb de RAM el algoritmo en el factor espacio consume 205,5 Mb.
Maquina 2: Serrano José	0,010	0,021	A razón de 12 Gb de RAM el algoritmo en el factor espacio consume 400 Mb.

Nota: La rapidez de ambos procesadores es constante por ende no influye en la implementación del algoritmo basado en reglas, por otro lado, se aprecia que el factor time medido en segundos depende del factor espacio que utiliza el algoritmo. La elaboración es propia y la fuente proporcionada por las investigaciones dadas.

Beneficiarios directos e indirectos del proyecto

Beneficiarios Directos

Se señalan como beneficiarios directos a las personas con trastorno dislábico que se sometan a un diagnóstico de déficit de fonemas basados en dislalia funcional. Adicional, las personas que presentan estos trastornos del lenguaje sabrán de manera directa el futuro tratamiento a seguir recomendado por el especialista logopédico.

Beneficiarios Indirectos

Se distinguen como beneficiarios indirectos a profesionales especialistas logopédicos, al padre de familia del paciente y a la comunidad, al adaptarse la parte tecnológica procediendo a la construcción de un algoritmo en ayuda del profesional para realizar diagnósticos basados en fichas logopédicas, además teniendo un algoritmo que pueda considerarse a futuro dentro del proyecto FCI de la facultad de Ciencias Matemáticas y Físicas, este último siendo otro beneficiario indirecto.

Entregables del proyecto

- ❖ Código fuente (Algoritmo) construido en la plataforma de Visual Code, en el respectivo lenguaje en este caso Python.
- ❖ Dataset con datos simulados utilizados para las respectivas pruebas.
- ❖ Artículo científico.

Propuesta

La propuesta tecnológica del presente proyecto de titulación se enfoca en desarrollar un prototipo de algoritmo para diagnóstico de trastorno dislábico en pacientes a partir de las fichas logopédicas.

Este algoritmo desarrollado tendrá como principal función arrojar un diagnóstico sobre cuáles son los fonemas con que presenta mayor complicación los pacientes y en qué posición se presenta dicha anomalía, esta puede ser al inicio, media o al final. Para poder llegar a este punto, el proceso que integra el algoritmo es que se tendrá una base de datos en Excel, donde se registrarán los déficit de fonemas del paciente, estarán clasificadas por 5 etapas como se indica en la siguiente tabla.

Tabla 39

Clasificación de déficit de fonemas.

Clasificación	Ponderación		
	Inicio	Medio	Final
Sustitución	1.1	1.2	1.3
Omisión	2.1	2.2	2.3
Distorsión	3.1	3.2	3.3
Inversión	4.1	4.2	4.3
Adición	5.1	5.2	5.3

Nota: La tabla indica los valores dentro de la clasificación del déficit de fonemas. La elaboración es propia y la fuente es proporcionada por las investigaciones realizadas.

Estos datos son leídos por el algoritmo y posteriormente son procesados, una vez culminado este paso, el algoritmo tiene integrado una funcionado denominada diagnosticar() en la cual dentro de ella se encuentran inmersas todas las reglas necesarias que aportaran al diagnóstico del paciente a evaluar, para ello el algoritmo recorre todo la data y va comparando los valores, los va clasificando y almacenando en los Array creados en el algoritmo, luego de haber recorrido toda la data del paciente, arrojará el diagnostico ya de forma ordenada, es decir en que fonemas presenta Sustitución, Omisión, Distorsión, Inversión y Adición y a su vez en qué posición también se diagnostica esta anomalía.

Con la realización de este proyecto se buscará obtener un diagnóstico en base a los resultados otorgados por el algoritmo, ya que en Ecuador existen muy pocos especialistas logopedas que ayuden a tratar este tipo de trastornos, por ello existe mayor demanda de pacientes a los que se les hace complicado asistir a estos centros médicos para que puedan tratar estos trastornos fonéticos. Los especialistas logopedas podrán dar mayor solución a los pacientes en el menor tiempo estimado y así poder erradicar y mejorar su calidad de vida, tanto en el ámbito educacional como social en cada uno de estos infantes.

Criterios de validación de la propuesta

Informe del plan de pruebas

En este punto, para poder validar nuestra propuesta planteada, se la realizara mediante pruebas inducidas las cuales tienen como características evaluar el desempeño del algoritmo y así conocer cuál es el tiempo que tardara en brindar un diagnóstico de los pacientes, cual es el porcentaje de precisión, para ello se realizaron las siguientes pruebas:

Prueba N°. 1: En la primera prueba realizada se ingresó un dataset que contiene 5 pacientes y a su vez presenta solo 6 fonemas a evaluar, en este caso se precisa que se emita el diagnóstico del paciente 1 y 3. Para la prueba del paciente 1 se indicó que todos los fonemas contengan un valor de 0, para así poder obtener como resultado que el infante no presenta problemas de dislalia en ninguno de los fonemas y por otra parte para el paciente 3 se fijaron diversos valores en cada fonema con la finalidad de obtener un diagnóstico con los tipos de trastornos de dislalia y en qué posición se encuentran , una vez realizada la prueba efectivamente los resultados fueron los pronosticados y el tiempo de ejecución para los dos casos fue de 0,03 seg., los casos exitosos obtenidos fueron 2, no se obtuvieron casos fallidos, no se presentaron complicaciones ni desviaciones cumpliendo así con la meta diaria de 2 casos en el día.

Prueba N°. 2: En esta segunda prueba realizada se ingresó un dataset que también contiene 5 pacientes y presenta 12 fonemas a evaluar, en este caso se precisa que se emita el diagnóstico del paciente 2 y 4. Para la prueba del paciente 2 se indicó que todos los fonemas contengan un valor de 0, para así poder obtener como resultado que el infante no presenta problemas de dislalia en ninguno de los fonemas y por otra parte para el paciente 4 se fijaron valores de 2.1 en cada fonema con la finalidad de obtener como diagnóstico que presenta problemas de Omisión en la posición inicial en todos los fonemas evaluados, una vez realizada la prueba efectivamente los resultados fueron los pronosticados y el tiempo de ejecución para los dos casos fue de 0,02 seg, los casos exitosos obtenidos fueron 2, no se obtuvieron casos fallidos, no se presentaron complicaciones ni desviaciones cumpliendo así con la meta diaria de 2 casos en el día.

Prueba N°. 3: Finalmente en esta tercera prueba realizada se ingresó un dataset que también contiene 5 pacientes y se ingresaron todos los 28 fonemas a evaluar, en este caso se

precisa que se emita el diagnóstico del paciente 5. Para la prueba del paciente se indicó que los primeros 14 fonemas contengan un valor de 1.1 y los otros 14 fonemas contengan un valor de 0, para así poder obtener como resultado que el infante presenta problemas de Sustitución en la posición inicial los primeros 14 fonemas, una vez realizada la prueba efectivamente el resultado fue el pronosticado y el tiempo de ejecución para este caso fue de 0,02 seg, los casos exitosos obtenidos fueron 1, no se obtuvieron casos fallidos, no se presentaron complicaciones ni desviaciones cumpliendo así con la meta diaria de 1 caso en el día.

Juicio de experto

Se tomará como criterio de validación de la propuesta el juicio de expertos los cuales serán los que ayudaran a validar la confiabilidad del presente proyecto, para ello se tomó en consideración a docentes de la carrera de Ingeniería en Sistemas de la Facultad de Ciencias Matemáticas y Físicas de la Universidad de Guayaquil y el criterio de una especialista en terapia del lenguaje del Centro de Rehabilitación SERLI, los cuales son: Ing. José Alonzo Anguizaca, Msc., Ing. Alfonso Guijarro Rodríguez, Ing., Charles Pérez Espinoza y Lcda. Diana Guerra Garnica respectivamente, véase en el Anexo 6: Anexo 6.1, Anexo 6.2, Anexo 6.3, Anexo 6.4 2

Para la realización del juicio de expertos se mostró todo el desarrollo y funcionamiento del algoritmo, explicando detalladamente cada una de sus partes, las clases y métodos que lo conforman, así mismo que detallo como se encuentra estructurado el dataset, los fonemas que contiene, el rango de valores que se definen en cada fonema y finalmente se puso a prueba el algoritmo para indicar que los resultados diagnósticos estén totalmente correctos y corroborar las validaciones de este. También se presenta una breve información de los expertos involucrados en la siguiente tabla.

Tabla 40*Información de Expertos*

Nombre y Apellidos	Profesión	Entidad	Años de Experiencia
José Alonzo Anguizaca	Ingeniero en Sistemas Computacionales	Universidad de Guayaquil	22 años
Alfonso Guijarro Rodríguez,	Ingeniero en Sistemas Computacionales	Universidad de Guayaquil	22 años
Charles Pérez Espinoza	Ingeniero en Sistemas Computacionales	Universidad de Guayaquil	6 años
Diana Guerra Garnica	Lcda. Terapia del lenguaje	Centro de Rehabilitación SERLI	6 años

Nota: Indica la información de los expertos involucrados, como sus nombres, la profesión que mantiene, entidad laboral actual y los años de experiencia. La elaboración es propia y la fuente obtenida a partir de las investigaciones realizadas.

Análisis de datos

En este proyecto de investigación, al contar con una encuesta formada por variables cualitativas se aplicarán las técnicas de tablas cruzadas, análisis de contingencia, estadístico Chi-Cuadrado para el contraste de hipótesis. Como se indica en la tabla 41 el resumen del procesamiento de los casos, de acuerdo a las preguntas cuatro y ocho.

Tabla 41

Resumen del procesamiento de los casos: Pregunta 4 y 8.

	Casos					
	Válidos		Perdidos		Total	
	N	Porcentaje	N	Porcentaje	N	Porcentaje
¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos? * ¿Se encuentra de acuerdo que al implementar un algoritmo para diagnóstico se llegara a beneficiar tanto al paciente como al especialista Logopeda?	63	100,0%	0	0,0%	63	100,0%

Nota: Resumen del procesamiento de los casos pregunta 4 y 8. La elaboración es propia y las fuentes obtenidas de las investigaciones realizadas.

Análisis: Las preguntas cuatro y ocho respectivamente son esenciales, porque la construcción de un algoritmo para diagnosticar el trastorno dislálico beneficiara directamente al paciente y al especialista logopeda respectivamente.

Análisis de contingencia

Las tablas de contingencia nos indica si estas interrelaciones son estadísticamente dependientes o en su defecto independientes, se efectúa el contraste estadístico Chi-cuadrado, prueba que se aplica para equiparar si ambas propiedades cualitativas permanecen en relación estadísticamente entre sí.

Tabla 42

¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos? Vs ¿Se encuentra de acuerdo que al implementar un algoritmo para diagnóstico se llegara a beneficiar tanto al paciente como al especialista Logopeda?

		¿Se encuentra de acuerdo que al implementar un algoritmo para diagnóstico se llegara a beneficiar tanto al paciente como al especialista Logopeda?				Total
		Totalmente en desacuerdo	Poco desacuerdo	Acuerdo	Muy de acuerdo	
¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos?	Si	Recuento 3	1	21	36	61
		% dentro de ¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos? 4,9%	1,6%	34,4%	59,0%	100,0%
¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos?	No	Recuento 0	1	1	0	2
		% dentro de ¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos? 0,0%	50,0%	50,0%	0,0%	100,0%
Total		Recuento 3	2	22	36	63
		% dentro del total 4,8%	3,2%	34,9%	57,1%	100,0%

Nota: Tabla de contingencia de la variable ¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos? y ¿Se encuentra de acuerdo que al implementar un algoritmo para diagnóstico se llegara a beneficiar tanto al paciente como al especialista Logopeda? La elaboración es propia y las fuentes obtenidas de las investigaciones realizadas.

Análisis: A raíz del análisis de la tabla de contingencia se aprecia que la relación de las variables cualitativas es favorable dado que existe un 59,0% de personas que, si consideran la construcción del algoritmo, y que están muy de acuerdo en que este beneficiara al paciente y al especialista logopeda, este porcentaje representa más de la mitad de los encuestados.

Chi cuadrado**Tabla 43***Pruebas de chi-cuadrado*

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	15,680 ^a	3	,001
Razón de verosimilitudes	6,827	3	,078
Asociación lineal por lineal	3,042	1	,081
N de casos válidos	63		

a. 6 casillas (75,0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es ,06.

Análisis: el nivel de significancia es (0,05) a su vez la significancia asintótica (bilateral) es 0,001 < 0,05 entonces si se acepta la hipótesis de que existe una relación entre ambos ítems (pregunta cuatro y ocho).

Estadísticos descriptivos

Tabla 44*Resumen del procesamiento de los casos*

		¿Tenía conocimiento acerca de este tema de la dislalia?	¿Conoce usted a familiares, amigos, infantes o personas cercanas que presenten alteración en alguno de estos fonemas?	¿Cree usted que tendría un impacto positivo adaptar en el área de la logopedia un algoritmo para diagnóstico?	¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos?	¿Qué tan importante considera que se genere una herramienta tecnológica para trastornos dislálicos?	¿Está de acuerdo usted en que se den a conocer estas herramientas tecnológicas en nuestra población?	¿Qué tan relevante cree usted que es dar una solución a estos problemas dislálicos?	¿Se encuentra de acuerdo que al implementar un algoritmo para diagnóstico se llegara a beneficiar al paciente como al especialista Logopeda?	¿Qué tan seguro se sentiría de usar esta herramienta, para que le genere un diagnóstico del trastorno dislálico?
N	Válidos	63	63	63	63	63	63	63	63	63
	Perdidos	0	0	0	0	0	0	0	0	0
Media		1,76	1,52	3,08	1,03	3,48	3,48	3,57	3,44	3,32
Mediana		2,00	2,00	3,00	1,00	4,00	4,00	4,00	4,00	3,00
Moda		2	2	3 ^a	1	4	4	4	4	3
Desv. típ.		,429	,503	1,036	,177	,564	,877	,530	,778	,591
Varianza		,184	,253	1,074	,031	,318	,770	,281	,606	,349
Asimetría		-1,260	-,098	-1,061	5,473	-,461	-1,850	-,630	-1,611	-,213
Error típ. de asimetría		,302	,302	,302	,302	,302	,302	,302	,302	,302
Curtosis		-,427	-2,057	,036	28,867	-,796	2,748	-,900	2,666	-,584
Error típ. de curtosis		,595	,595	,595	,595	,595	,595	,595	,595	,595
Rango		1	1	3	1	2	3	2	3	2
Mínimo		1	1	1	1	2	1	2	1	2
Máximo		2	2	4	2	4	4	4	4	4
	25	2,00	1,00	3,00	1,00	3,00	3,00	3,00	3,00	3,00
Percentiles	50	2,00	2,00	3,00	1,00	4,00	4,00	4,00	4,00	3,00
	75	2,00	2,00	4,00	1,00	4,00	4,00	4,00	4,00	4,00

Nota: Tabla de estadísticas descriptivas en base a las preguntas de la encuesta. La elaboración es propia y las fuentes obtenidas de las investigaciones realizadas.

Análisis : Como se aprecia en la tabla 43 los estadísticos de tendencia central (media, mediana y moda) obtuvieron puntuaciones positivas dentro de la encuesta, a excepción de la primera y segunda pregunta de la encuesta, las cuales evaluaban el conocimiento sobre la dislalia como enfermedad, por el contrario las siguientes preguntas nos ayudaron a contrarrestar este trastorno dislítico con lo que estos resultados aprobaron la construcción del algoritmo para diagnóstico, a su vez considerando los percentiles que presentan la similitud necesaria al estadístico de tendencia central mediana, los datos son estadísticamente satisfactorios, por otra parte los estadísticos de forma (asimetría, curtosis) se interpretan que indican el valor intermedio como duda a contestar las preguntas considerando que en su mayoría las respuestas eran positivas.

Resultados

El algoritmo basado en reglas y el de aprendizaje supervisado Naive Bayes fueron empleado con el uso de la herramienta tecnológica Python, como es de conocimiento general este lenguaje de programación es uno de los más utilizados a nivel mundial y se adapta en gran parte a los requerimientos que se lo aplique. La codificación de estos algoritmos fue desarrollada bajo el entorno Visual Studio Code, cuya finalidad es la obtención del diagnóstico de trastornos dislíticos a partir de la información ingresada por el dataset.

Como primer resultado se obtuvo que con la aplicación del algoritmo Naive Bayes no se otorgaba un diagnóstico esperado, por tal motivo no fue recomendable su uso. Por otra parte, con el algoritmo basado en reglas se obtuvo un éxito bastante favorable en comparación al algoritmo anterior, dándole así mayor importancia y finalmente siendo el más indicado para su desarrollo y

aplicación el cual cumple con brindar un diagnóstico, acotándole que los tiempos de procesamiento también son bastante cortos lo cual lo hace que sea un algoritmo eficiente.

- **Objetivo 1:** Con la ayuda de las fuentes de información científica se pudo obtener toda la información más relevante que aportaron a definir los tipos de trastornos dislálicos y conocer cuáles son sus principales causas, para que posteriormente con todo lo investigado poder estructurar de manera adecuada el algoritmo basado en reglas utilizado.
- **Objetivo 2:** Se obtuvo una base de datos por parte del tutor con datos reales los cuales nos sirvieron como guía para posteriormente también poder realizar una simulación y depuración de datos haciendo uso de la herramienta tecnológica @risk 8.3.
- **Objetivo 3:** Finalmente teniendo bien estructurado el dataset, se lo puso a prueba dentro del algoritmo basado en reglas, en el cual se obtuvieron los resultados esperados, cumpliendo con las expectativas propuestas y dando por afirmado que el diagnostico resultante es correcto.

CAPÍTULO IV

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

- Se realizaron las investigaciones sobre las diferentes pruebas de evaluación denominadas fichas logopédicas, analizando fuentes bibliográficas, artículos científicos, estudio de casos similares con las cuales se pudo obtener los argumentos válidos para la construcción del algoritmo de diagnóstico.
- Se analizó y según la comparativa de los diferentes tipos de algoritmos investigados como lo son: basado en reglas, aprendizaje supervisado como Naive Bayes, K-NN, Árboles de decisiones, se pudo concluir mediante indicadores cuál de estos algoritmos cuenta con mayor eficiencia, arrojando como resultado que el óptimo al momento de generar el diagnóstico es el basado en reglas en relación con la base de datos utilizada.
- Se construyó un algoritmo el cual será el encargado de generar un diagnóstico de trastornos dislálicos en cada uno de los pacientes a evaluarse, a su vez se realizó una simulación de los datos utilizando herramientas tecnológicas tales como @Risk 8.3, Stat::Fit para ser utilizados dentro del presente proyecto
- Finalmente, se redactó el artículo científico que contiene todas las etapas de la investigación de este trabajo dejando así una memoria informativa, con el cual se pretende ayudar a futuros trabajos pues se menciona el desarrollo del algoritmo basado en reglas y los casos de trastornos fonológicos que comprende la dislalia funcional.

Recomendaciones

- Se recomienda implementar algoritmos de Machine Learning cómo por ej.: Redes neuronales, aprendizaje automático, algoritmos genéticos, entre otros, para procesar la información con mayor precisión y obtener resultados más favorables al estudio.
- Se recomienda complementar el proyecto con una interfaz y agregar un módulo de control y seguimiento para controlar el estado de los pacientes durante el tratamiento e incluso si tiene algún inconveniente cambiarle el tratamiento de tal manera que el paciente se sienta a gusto y pueda aprender de una mejor manera los fonemas.

Trabajos futuros

Entre los trabajos futuros se puede destacar los siguientes:

- Se podría utilizar este proyecto para ayudar a otros tipos de trastornos cómo por Ej.: afasia, autismo, dislexia, entre otros, de esa manera poder dar terapias que los ayude con sus distintos trastornos y puedan tener una mejor calidad de vida.
- También, se puede implementar una página web o aplicación móvil interactiva que, por medio de imágenes, audios y videos, inciten a los pacientes a hablar a través de un micrófono y dependiendo de su desenvolvimiento detecte el porcentaje de equivocación, las causas del error y diagnosticar el nivel de dislalia que presente dicho paciente e incluso pronosticar si se puede corregir a tiempo o no.

REFERENCIAS BIBLIOGRÁFICAS

- D'Angiolo , F. G., Kwist, I. F., Loiseau, M., Contreras, D. E., & Asteasuain, F. (Octubre de 2019). *Repositorio Institucional de la UNLP*. Obtenido de Algoritmos de regresión lineal aplicados al mantenimiento de un datacenter: <http://sedici.unlp.edu.ar/handle/10915/91401>
- Alarcón, D., Marchi, J., Hernández , J., Chávez, O., & Hernández Ocaña, B. (2020). Clasificadores basados en reglas y selección de atributos para el diagnóstico clínico de subtipos del Síndrome de Guillain-Barré. *rsc.cic.ipn.mx*, 1-13. Obtenido de https://rsc.cic.ipn.mx/2020_149_8/Clasificadores%20basados%20en%20reglas%20y%20seleccion%20de%20atributos%20para%20el%20diagnostico%20clinico.pdf
- ALBOR-COHS, G. (1999). *ELA-R: Examen logopedico de articulación (revisado)*. Obtenido de ACADEMIA:
https://www.academia.edu/8875885/ELA_Manual_de_Referencia_ALBOR_COHS
- Ali, M., Jung, L. T., Abdel-Aty, A.-H., Abubakar, M. Y., Elhoseny, M., & Ali, I. (2020). Semantic-k-NN algorithm: An enhanced version of traditional k-NN algorithm. *Expert Systems with Applications vol.151*, 9, doi: <https://doi.org/10.1016/j.eswa.2020.113374>.
- Antúñez, C. (2017). *Iniciación a excel 2013*. Editorial Elearning, SL.
- aprendemachinelearning. (10 de Julio de 2018). *Clasificar con K-Nearest-Neighbor ejemplo en Python*. Obtenido de KNN: <https://www.aprendemachinelearning.com/clasificar-con-k-nearest-neighbor-ejemplo-en-python/>
- Ardila, E. (2020). COVID-19 en Colombia e inmunidad de rebaño: ¿es momento de considerarla? *Revista Colombiana de Endocrinología, Diabetes y Metabolismo*, 57-59.

- Arenal Laza, C. (2019). *Investigación y recogida de información de mercados*. Tutor Formación. Obtenido de https://books.google.com.ec/books?id=srenDwAAQBAJ&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false.
- Aquilla Morocho, E. F. (2021). *MINERÍA DE OPINIÓN PARA TEXTOS EN ESPAÑOL USANDO PROCESAMIENTO NATURAL DEL LENGUAJE*. Quito: Universidad Politécnica Salesiana. <http://dspace.ups.edu.ec/handle/123456789/20696>.
- Barboza, A., & otros, &. (2020). Opinión de la población migrante sobre los cuidados del equipo de salud en dos policlínicas municipales de Montevideo. *Scielo*, 44-53.
- Botella, J., & Zamora, A. (2017). El meta-análisis: una metodología para la investigación en educación. *Educación XXI*, 20(2), 17-38, doi: 10.5944/educXX1.18241.
- Burgo, O. B., León, J. L., Cáceres, M. L., Pérez, C. J., & Espinoza, E. E. (2019). Some thoughts on research and educational intervention. *Revista Cubana de Medicina Militar*, 48(Supl. 1), e383, Recuperado en 29 de julio de 2021, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0138-65572019000500003&lng=es&tlng=en.
- Cabrera Cárdenas, A. M., & Jiménez Cabrera, K. E. (2017). *PREVALENCIA DE FONEMAS ALTERADOS EN NIÑOS Y NIÑAS DEL DISTRITO*. Obtenido de Repositorio Institucional Universidad de Cuenca: <https://dspace.ucuenca.edu.ec/bitstream/123456789/28760/1/PROYECTO%20DE%20INVESTIGACION.pdf>

- Cardoso, C., Talame, L., & Amor, M. (2021). Aprendizaje automático aplicado a la pandemia del virus Covid-19 en Argentina ESPERADOS. *XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja)*, 78-81.
- Carrasco-Machado, J., Ramos, A. M., & Morales-Muñoz, R. (2020). Un modelo probabilístico de inventarios por demanda independiente mediante el uso de GRASP. *Ecuadorian Science Journal*, 4(1), 26-31.
- Caturla, R., & Congosto, C. (2016). *Sistema de creación de tutoriales interactivos para el aprendizaje de lenguajes de programación*. Universidad Complutense de Madrid .<https://eprints.ucm.es/id/eprint/38408>.
- Cedeño Pinargote, K. E., & Chinga Quimis, E. A. (2020). *Métricas y pesos para el establecimiento de similitudes en rasgos para el sistema de razonamiento basado en casos, del proyecto FCI para pacientes con Dislalia*. Obtenido de Repositorio Universidad de Guayaquil: <http://repositorio.ug.edu.ec/bitstream/redug/49648/1/B-CISC-PTG-1843-2020%20Cede%c3%b1o%20Pinargote%20Katherine%20Edith%20-%20Chinga%20Quimis%20Edisson%20Alberto.pdf>
- Chan May, O. A., Peña Koo, J. J., Vianne Kinani, J. M., & Zapata Encalada, M. A. (2018). Construcción De Un Modelo De Predicción Para Apoyo Al Diagnóstico De Diabetes (Construction of a Prediction Model To Support the Diabetes Diagnosis). *Pistas Educativas*, 2105-2122.
- Chaparro Pesca, J. A., & Cepeda Araque , C. H. (Agosto de 2018). *Repositorio UPTC*. Recuperado el 2021, de Aplicación de un Modelo de Regresión Logística con Respuesta Politémica Nominal en el Análisis de Preferencias Alimentarias de Aves:

http://rdigitales.uptc.edu.co/memorias/index.php/apli_estad/apli_esta2016/paper/view/21

76

Chen, J. (2019). A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Elsevier - Environment International*, 1-14.

Cobo Viera, R. E., & Jaramillo Armijos, B. V. (Febrero de 2020). *Construcción de una base de casos y base transaccional para el sistema del Proyecto FCI Herramienta para el tratamiento de trastornos fonéticos (Dislalia)*. Guayaquil: Repositorio de la Universidad de Guayaquil. <http://repositorio.ug.edu.ec/handle/redug/48980>. Recuperado el Agosto de 2021

Davidson, L., & González Ibáñez, Á. (2017). La recuperación centrada en la persona y sus implicaciones en salud mental. *Revista de la Asociación Española de Neuropsiquiatría*, 189-205.

Del Valle Benavides, A. R. (2017). Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones. *Departamento de Estadística e Investigación Operativa*.

ERP. (2021). *Análisis de requisitos técnicos: procesos y procedimientos*. Obtenido de EVALUANDO ERP.COM: <https://www.evaluandoerp.com/software-erp/conceptos-erp/analisis-requisitos-tecnicos/>

Escobar Gallegas, J. L., & Bilbao Ramirez, P. H. (2020). *INVESTIGACION Y EDUCACION SUPERIOR*. LULU COM.

EspacioLogopedico. (2021). *ELCE: Exploracion del lenguaje comprensivo y expresivo*.

Obtenido

de

EspacioLogopedico.com:

<https://www.espaciologopedico.com/tienda/prod/576/elce-exploracion-del-lenguaje-comprensivo-y-expresivo.html>

Espinoza, C. (2018). *Diseño y desarrollo de una herramienta lúdica informática para estimular e intervenir los niveles de lenguaje en niños desde los 4 a 5 años*. Cuenca: <https://dspace.ups.edu.ec/>.

Flores Rodríguez, C. (Marzo de 2021). *Repositorio Institucional Universidad Veracruzana*. Veracruz: Universidad Veracruzana. <https://www.google.com/search?client=firefox-b-d&q=Generaci+%CC%81on+de+%CC%81arboles+dedecisi+%CC%81on+usando+unalgoritmo+inspirado+en+laF+%CC%81%C4%B1sica#>. Recuperado el Julio de 2021, de Generación de árboles de decisión usando un algoritmo inspirado en la Física: <https://www.google.com/search?client=firefox-b-d&q=Generaci+%CC%81on+de+%CC%81arboles+dedecisi+%CC%81on+usando+unalgoritmo+inspirado+en+laF+%CC%81%C4%B1sica#>

Frascaroli, M. (6 de Junio de 2020). *Facilidad de gestión y rápida implementación: claves en la elección de soluciones de servicio al cliente*. Obtenido de Casos y entrevistas, Chatbots: <https://planetachatbot.com/facilidad-gestion-y-rapida-implementacion-claves-en-eleccion-soluciones-cliente/>

Fredes Albarracín, E. (2019). *Dislalia: Tipos y tratamientos*. Obtenido de OirPensarHablar: <https://oirpensarhablar.com/dislalia-tipos-tratamiento/>

Fuentes, M. D., & Medina, W. D. (Abril de 2021). *Diseño de un modelo predictivo-asistencial de pacientes infectados por Covid-19, mediante un modelo supervisado de machine learning basado en criterio de derivación hospitalaria o ambulatoria*. Obtenido de Repositorio UG: <http://repositorio.ug.edu.ec/bitstream/redug/52650/1/B-CISC-PTG->

1901-2021%20Fuentes%20Marmolejo%20Melina%20Daniela%20-%20Medina%20Parra%20Wilmer%20David.pdf

Galeano, M. E. (2020). Parte 3. Modalidades de investigación social cualitativa. En M. Galeano, *Diseño de proyectos en la investigación cualitativa* (págs. 13-89). Recuperado de https://books.google.com.ec/books?id=Xkb78OSRMI8C&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false.

Galvez Quilca, R. (2019). *SISTEMA WEB PARA EL PROCESO DE CONTROL DE PROYECTOS EN LA EMPRESA HUNDRED S.A.C.* César Vallejo: Repositorio Digital Institucional Universidad César Vallejo. <https://hdl.handle.net/20.500.12692/63935>.
Obtenido de Repositorio Digital Institucional Universidad César Vallejo.

Galvez Quilca, R. (2019). *SISTEMA WEB PARA EL PROCESO DE CONTROL DE PROYECTOS EN LA EMPRESA HUNDRED S.A.C.* Obtenido de Repositorio Digital Institucional Universidad César Vallejo: <https://hdl.handle.net/20.500.12692/63935>

García Girón, E. G. (2020). *Detección de patrones de deserción estudiantil mediante aplicación de Árboles de Decisión C4. 5 en el IESTP “Señor de Chocán” de Querecotillo.* Universidad Cesar Vallejo. <https://hdl.handle.net/20.500.12692/55029>.

García Torres, M. E., & Espino Quiñones, L. (2018). *Aplicación de minería de datos basado en árboles de decisión para predecir el riesgo de morosidad de los clientes en la empresa de seguros Oncosalud SAC 2018.* Lima: Universidad Autónoma del Perú. <http://repositorio.autonoma.edu.pe/handle/AUTONOMA/700>.

Geer Mountain Software Corp. (2016). *Stat::Fit*. Obtenido de <https://www.geerms.com/files/114225421.pdf>

Gilbets, L. (11 de Noviembre de 2020). *Qué es la metodología Kanban y cómo utilizarla*.

Obtenido de IEBS: <https://www.iebschool.com/blog/metodologia-kanban-agile-scrum/>

González González, L. (2005). *Análisis clínico de la comunicación Verbal en niños con dificultades para aprender que cursan el segundo grado en la enseñanza especial*.

Obtenido de Repositorio Institucional de la Universidad Central "Marta Abreu" de Las Villas (UCLV):

<https://dspace.uclv.edu.cu/bitstream/handle/123456789/4252/Liliana%20Teresita%20Glez%20viejas.pdf?sequence=1&isAllowed=y#:~:text=La%20investigaci%C3%B3n%20del%20lenguaje%20impresivo,1%C3%B3gico%20gramaticales%20y%20fragmentos%20literarios.>

Guevara Alban, G. P., Verdesoto Arguello, A. E., & Castro Molina, N. E. (2020). Metodologías de investigación educativa (descriptivas, experimentales, participativas, y de investigación-acción). *Revista científica Mundo de la Investigación y el conocimiento*, 163-173.

Guillén Escamilla, J. E. (Nov de 2020). *Algunas consideraciones sobre el panorama actual de la lingüística clínica*. Obtenido de *Literatura y lingüística* (42) 303-329: <http://dx.doi.org/10.29344/0717621x.42.2598>

Guzmán Rizo, A. (Mayo de 2020). *Repositorio Institucional de la Universidad de Costa Rica*. Recuperado el Julio de 2021, de Comparación de dos estrategias metodológicas para la estratificación socioeconómica del marco muestral de viviendas de Costa Rica: <http://www.kerwa.ucr.ac.cr/handle/10669/81098?locale-attribute=en>

Hernández , C. E., & Carpio, N. (2019). Introducción a los tipos de muestreo. *Revista científica del Instituto Nacional de Salud*, 76-79.

- Hernández Escobar, A. A., Ramos Rodríguez, M. P., Placencia López, B. M., Indacochea Ganchozo, B., Quimis Gómez, A. J., & Moreno Ponce, L. A. (2018). *METODOLOGIA DE LA INVESTIGACION CIENTIFICA*. Editorial Área de innovación y desarrollo.
- Hernández Escobar, A. A., Ramos Rodríguez, M. P., Placencia López, B. M., Indacochea Ganchozo, B., Quimis Gómez, A. J., & Moreno Ponce, L. A. (2018). Metodología de la investigación científica. En A. Hernández Escobar, *Capítulo 2: La Ingeniería ambiental como ciencia* (págs. 15-174). doi : <http://dx.doi.org/10.17993/CcyL1.2018.15>.
- INEC. (28 de Noviembre de 2010). *Resultados Provinciales*. Obtenido de INEC: <https://www.ecuadorencifras.gob.ec/wp-content/descargas/Manu-lateral/Resultados-provinciales/guayas.pdf>
- Leon Atiquipa, H. E. (Abril de 2018). *Repositorio Institucional de la PUCP*. Recuperado el 2021, de Desarrollo de un modelo algorítmico basado en árboles de decisión para la predicción de la permanencia de un paciente en un proceso psicoterapéutico: <http://repositorio.pucp.edu.pe/index/handle/123456789/149024>
- Lizares Castillo, M. (2017). Comparación de modelos de clasificación: regresión logística y arboles de clasificación para evaluar el rendimiento académico. *UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS*, 1-68.
- Logopedia. (21 de diciembre de 2019). *¿Qué es un Logopeda?* Obtenido de Logopedia y más: <https://www.logopediaymas.es/blog/que-es-un-logopeda/>
- López Herrero, R. (2020). *Propuesta de intervención en un niño con dislalias y TDHA*. Obtenido de Repositorio Documental Universidad de Valladolid: <https://uvadoc.uva.es/bitstream/handle/10324/42960/TFG-G4372.pdf?sequence=1&isAllowed=y>

- Lynn Speiser , J., E. Miller, M., Tooze, J., & Lp, E. (November de 2019). A comparison of random forest variable selection methods for classification prediction modeling. *Science Direct*, 93-101. Recuperado el 2021, de A comparison of random forest variable selection methods for classification prediction modeling: <https://www.sciencedirect.com/science/article/abs/pii/S0957417419303574>
- Maillo, J., Luengo, J., Garcia, S., Herrera, F., & Triguero, I. (2018). *Un enfoque aproximado para acelerar el algoritmo de clasificacion Fuzzy kNN para Big Data*. Granada: Universidad de Granada, España. https://sci2s.ugr.es/sites/default/files/ficherosPublicaciones/2514_CAEPiA2018_paper_184.pdf.
- Maisueche Cuadrado, A. (2019). *UTILIZACIÓN DEL MACHINE LEARNING EN LA INDUSTRIA 4.0*. Valladolid: UNIVERSIDAD DE VALLADOLID. <https://core.ac.uk/download/pdf/228074134.pdf>.
- Manrique Rojas, E. (2020). Machine Learning: análisis de lenguajes y herramientas para desarrollo. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 586-599.
- Márquez, F., Ginberg, M., Rivas, M. E., Jarmoluk, V., Bonamico, L., Ron, M., . . . Olmos, L. (2020). Aplicación del aprendizaje automático para predecir la recuperación de pacientes con desordenes prolongados de conciencia. *Congreso Argentino de Informática y Salud*, 228-247.
- Martínez Manzo, R. (2020). *SISTEMA DE MENÚ INTERACTIVO APOYADO POR UN ALGORITMO BASADO EN REGLAS DE ASOCIACIÓN PARA LA ALIMENTACIÓN DE PACIENTES INTERNOS EN UN HOSPITAL. CASO DE ESTUDIO: HOSPITAL DEL*

PACÍFICO. ACAPULCO DE JUAREZ: TECNOLÓGICO NACIONAL DE MÉXICO.

<https://rinacional.tecnm.mx/jspui/handle/TecNM/1465>.

Martínez R., J. L. (2019). El proceso de elaboración y validación de un instrumento de medición documental. *Revista Especializada Acción y Reflexión Educativa*, 2.

Menacho Chiok, C. H. (2017). Predicción del rendimiento académico aplicando técnicas de minería de datos. *Anales Científicos*, 26-33.

Merino, J. (2017). La potencialidad de la Regresión Logística Multinivel. *Empiria. Revista De metodología De Ciencias Sociales*, (36), 177-211.
<https://doi.org/10.5944/empiria.36.2017.17865>.

Microsoft. (8 de Julio de 2018). *Algoritmo de árboles de decisión de Microsoft*. Obtenido de Microsoft: <https://docs.microsoft.com/es-es/analysis-services/data-mining/microsoft-decision-trees-algorithm?view=asallproducts-allversions>

Microsoft. (2021). *Actualizaciones Visual Studio Code*. Obtenido de Visual Studio Code : https://code.visualstudio.com/updates/v1_60

Monja Sandoval, E. A. (2020). *Comparación de naive bayes y series de tiempo en la predicción de morosidad de cuotas sociales del Colegio de Ingenieros del Perú Consejo Departamental Lambayeque*. Pimentel: Universidad Señor de Sipán.
<https://hdl.handle.net/20.500.12802/6755>.

Montoya Zuluaga, C. A. (Julio de 2020). *Competencias ciudadanas para jóvenes en inmersión política*. Pereira: Universidad de Pereira.
<https://repositorio.ucp.edu.co/bitstream/10785/6343/1/DDEPDH93.pdf>. Recuperado el Agosto de 2021

- Olmedo Vega, V., Aguilar Idáñez, J., & Arenillas Lara, J. F. (2019). ANÁLISIS DE FACTORES ASOCIADOS A LA RECUPERACIÓN INTEGRAL DE PACIENTES DE ICTUS AL ALTA DE HOSPITAL DE AGUDOS. *Revista Española de Salud Pública*, 1-13.
- Origel, C. G., Rendón , E., Abundez , I., & Alejo, R. (2020). *Repositorio Digital Instituto Politécnico Nacional*. Recuperado el 2021, de Redes neuronales artificiales y árboles de decisión para la clasificación con datos categóricos: https://rcs.cic.ipn.mx/2020_149_8/Redes%20neuronales%20artificiales%20y%20arboles%20de%20decision%20para%20la%20clasificacion%20con%20datos%20categoricos.pdf
- Otzen , T., & Manterola, C. (2017). Técnicas de Muestreo sobre una Población a Estudio. *Scielo*, 227-232.
- Pascual Garcia, P. (2012). Parte III. Conceptos y clasificación de las dislalias. En P. Pascual, *La dislalia: naturaleza, diagnóstico y rehabilitación* (pág. 194. Recuperado de <https://elibro.net/es/ereader/uGuayaquil/153571>). España: CEPE.
- Pérez Alonso, D. (Febrero de 2017). *Utilización de Técnicas de Minería de Dato para la Predicción del Comportación de Biosólidos Aplicados en Agricultura*. Burgos: UNIVERSIDAD DE BURGOS. doi: <http://hdl.handle.net/10259/4509>. Recuperado el 2021, de Utilización de Técnicas de Minería de Dato para la Predicción del Comportación de Biosólidos Aplicados en Agricultura.
- Pérez Díaz, K. N. (2021). PREDICCIÓN DE USO DE UNIDADES DE CUIDADOS INTENSIVOS PARA LA PANDEMIA COVID-19. *UNIVERSIDAD DE CONCEPCIÓN*, 1-87.

Pérez, M. (11 de febrero de 2021). *Definición de Trastorno*. Obtenido de ConceptoDefinición:
<https://conceptodefinicion.de/trastorno/>

Python. (2020). *Python*. Obtenido de Python Software Foundation:
<https://www.python.org/about/>

Quan, Z., & Valdez, E. A. (July de 2018). *De Gruyter Publishes* . Recuperado el 2021, de
 Predictive analytics of insurance claims using multivariate decision trees:
<https://www.degruyter.com/document/doi/10.1515/demo-2018-0022/html>

RAE. (2020). Obtenido de Diccionario de la lengua española: <https://dle.rae.es/>

Raffino, M. (11 de noviembre de 2020). *Fonema*. Obtenido de Concepto.de:
<https://concepto.de/fonema/>

Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje
 supervisado aplicando regresión logística. (2020). : *XXVI Congreso Argentino de
 Ciencias de la Computación*, 491-500.

Regalado Macias, R. M., & Calle De la A, J. J. (2020). *Construcción de algoritmo de similitud
 para el sistema del Proyecto FCI, Herramienta para el tratamiento de trastornos
 fonéticos (Dislalia)*. Obtenido de Repositorio UG:
[http://repositorio.ug.edu.ec/bitstream/redug/48877/1/B-CISC-PTG-1728%20Regalado%20Mac%
 c3%adas%20Rosa%20Mar%c3%ada%20-%20Calle%20de%20la%20A%20%20Juan%20Jos%c3%a9.pdf](http://repositorio.ug.edu.ec/bitstream/redug/48877/1/B-CISC-PTG-1728%20Regalado%20Mac%c3%adas%20Rosa%20Mar%c3%ada%20-%20Calle%20de%20la%20A%20%20Juan%20Jos%c3%a9.pdf)

Regalado Macias, R. M., & Calle De la A, J. J. (2020). *Construcción de algoritmo de similitud
 para el sistema del Proyecto FCI, Herramienta para el tratamiento de trastornos
 fonéticos (Dislalia)*. Guayaquil: Universidad de Guayaquil.
<http://repositorio.ug.edu.ec/bitstream/redug/48877/1/B-CISC-PTG->

1728%20Regalado%20Mac%c3%adas%20Rosa%20Mar%c3%ada%20-

%20Calle%20de%20la%20A%20%20Juan%20Jos%c3%a9.pdf. Obtenido de Repositorio UG.

Reyes González, Y., Martínez Sánchez, N., Díaz Sardiñas, A., & Milanés Luque, M. (2017). De la extracción al modelado del conocimiento en un Sistema Basado en el Conocimiento. Un enfoque desde el agrupamiento conceptual lógico combinatorio (From the extraction to knowledge modeling in a Knowledge Based System). . *GEOCONTEC: Revista Internacional De Gestión Del Conocimiento Y La Tecnología*, 5 (2), 41-57 Recuperado a partir de: <https://www.upo.es/revistas/index.php/gecontec/article/view/2549>.

Ríos Sucasaca, K. D. (2017). REGRESIÓN LOGÍSTICA DE LOS FACTORES QUE INFLUYEN EN LA GASTRITIS, HOSPITAL REGIONAL MANUEL NUÑEZ BUITRÓN-PUNO. *UNIVERSIDAD NACIONAL DEL ALTIPLANO*, 1-139.

Rivas, R. (11 de julio de 2021). *APOYO ACADÉMICO POR ANTOLOGÍAS*. Obtenido de UNAM: <https://www.iztacala.unam.mx/rrivas/diagnostico.html>

Robles Pastor, B. F. (Febrero de 2019). *Pueblo Continente* . Recuperado el Julio de 2021, de Población o muestra: <http://200.62.226.189/PuebloContinente/article/view/1269/1099>

Rodríguez, M., & Mendivelso, F. (2018). DISEÑO DE INVESTIGACIÓN DE CORTE TRANSVERSAL. *unisanitas.edu.co*, 1-7.

Roman, V. (25 de Abril de 2019). *Algoritmos Naive Bayes: Fundamentos e Implementación*. Obtenido de Ciencia & Datos: <https://medium.com/datos-y-ciencia/algoritmos-naive-bayes-fudamentos-e-implementaci%C3%B3n-4bcb24b307f>

- Romero Tabeayo, I., & Rosa N., F. C. (2017). Diseño de un instrumento de observación para la investigación en enseñanza . *REIPE (Revista de Estudios e Investigación en Psicología y Educación)*, 170-173 .
- Rouhiainen, L. (2018). *Inteligencia artificial 101 cosas que debes saber hoy sobre nuestro futuro*. Madrid: Alienta Editorial .
- Sagaró del Campo, N. M., & Zamora Matamoros, L. (2019). Análisis estadístico implicative versus Regresión logística binaria para el estudio de la causalidad en salud. *Scielo*, 1416-1440.
- Salazar P, C., & Del Castillo G, S. (2018). *Fundamentos básicos de estadística*.
- Salazar, E. J., & Alzate, W. A. (2018). Aplicación de la simulación Monte Carlo en la proyección del estado de resultados. Un estudio de caso. *Espacios 39(51)*, 11 . Recuperado de <http://es.revistaespacios.com/a18v39n51/a18v39n51p11.pdf>.
- Salazar, E. J., & Alzate, W. A. (2018). Aplicación de la simulación Monte Carlo en la proyección del estado de resultados. Un estudio de caso. *Espacios 39(51)*, 11 . Recuperado de <http://es.revistaespacios.com/a18v39n51/a18v39n51p11.pdf>.
- Salesforce Latinoamérica. (22 de junio de 2017). *Inteligencia Artificial: ¿Qué es? - Blog de Salesforce*. Obtenido de [salesforce blog: https://www.salesforce.com/mx/blog/2017/6/Que-es-la-inteligencia-artificial.html](https://www.salesforce.com/mx/blog/2017/6/Que-es-la-inteligencia-artificial.html)
- Sánchez Pacheco, C. L. (2019). Gamificación: Un nuevo enfoque para la educación ecuatoriana Gamification: A new approach to Ecuadorian education. *Revista Internacional Docente 2.0 Tecnología-Educativa*, 1-10.
- Sánchez Terán, M. d. (2018). *GUÍA DIDÁCTICA DE DETECCIÓN E INTERVENCIÓN DE LA DISLALIA FUNCIONAL EN NIÑOS DE 4 A 6 AÑOS*. Obtenido de Repositorio de la

Pontifica Universidad Católica del Ecuador:

<http://repositorio.puce.edu.ec/bitstream/handle/22000/15119/Tesis->

[Mari%CC%81a%20del%20Carmen%20Sa%CC%81nchez%20Tera%CC%81n%20.pdf?](http://repositorio.puce.edu.ec/bitstream/handle/22000/15119/Tesis-Mari%CC%81a%20del%20Carmen%20Sa%CC%81nchez%20Tera%CC%81n%20.pdf?sequence=1&isAllowed=y)

[sequence=1&isAllowed=y](http://repositorio.puce.edu.ec/bitstream/handle/22000/15119/Tesis-Mari%CC%81a%20del%20Carmen%20Sa%CC%81nchez%20Tera%CC%81n%20.pdf?sequence=1&isAllowed=y)

Sandoval, L. J. (2018). Algoritmos de Aprendizaje Automático para Análisis y Predicción de Datos. *Revista Tecnológica ITCA-FEPADE*, 36-40.

Schwarz Díaz, M. (2018). *Identificación y caracterización del problema de investigación para la elaboración de la tesis universitaria*. Lima: Repositorio de la Universidad de Lima.

[https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj6xqyRkq3yAhWhSjABHXQqA6UQFnoECAwQAw&url=https%3A)

[8&ved=2ahUKEwj6xqyRkq3yAhWhSjABHXQqA6UQFnoECAwQAw&url=https%3A](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj6xqyRkq3yAhWhSjABHXQqA6UQFnoECAwQAw&url=https%3A)

[%2F%2Frepositorio.ulima.edu.pe%2Fhandle%2F20.500.12724%2F7099&usg=AOvVaw](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj6xqyRkq3yAhWhSjABHXQqA6UQFnoECAwQAw&url=https%3A)

[1QhzkA](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj6xqyRkq3yAhWhSjABHXQqA6UQFnoECAwQAw&url=https%3A). Recuperado el 2021

Shaier, S. (19 de Febrero de 2019). *ML Algorithms: One SD (σ)- Bayesian Algorithms*. Obtenido de Towards data science: <https://towardsdatascience.com/>

Shaier, S. (19 de Febrero de 2019). *ML Algorithms: One SD (σ)- Bayesian Algorithms*. Obtenido de Towards data science: <https://towardsdatascience.com/>

Significados. (11 de Enero de 2017). *Significado de Excel*. Obtenido de Significados: <https://www.significados.com/excel/>

Solíz Plata, D. J. (2019). *Cómo Hacer Un Perfil Proyecto De Investigación Científica*. Palibrio

Recuperado de [https://books.google.com.ec/books?id=Q-](https://books.google.com.ec/books?id=Q-GCDwAAQBAJ&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false)

[GCDwAAQBAJ&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false](https://books.google.com.ec/books?id=Q-GCDwAAQBAJ&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false).

- Soni, D. (8 de Junio de 2018). *Introduction to Bayesian Networks*. Obtenido de Toward data science: <https://towardsdatascience.com/introduction-to-bayesian-networks-81031eed94e>
- Sznajdleder, P. (2017). *Algoritmos a fondo: con implementaciones en c y java*. Argentina: Alfaomega.
- Vega Hernández, C. A. (2020). ESTRATIFICACIÓN DE RIESGO DE INFECCIONES BACTERIANAS INVASORAS BASADO EN ALGORITMOS DE MACHINE LEARNING PARA PACIENTES PEDIÁTRICOS INGRESADOS POR NEUTROPENIA FEBRIL. *UNIVERSIDAD DE CHILE FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS DEPARTAMENTO DE INGENIERÍA INDUSTRIAL*, 1-99.
- Vicuña Pinto, A. E., Cortez Chichande , B. S., & Basurto Tovar , Y. J. (2018). Determinación de modelos predictivos para los indicadores de competitividad empresarial aplicando regresión lineal. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 94-107.
- Villabona, J. (4 de Abril de 2018). Base de conocimientos. *¿Cuáles son los requerimientos mínimos que se necesitan en un computador para trabajar exitosamente en el aula virtual?*
- Walker, J. S. (2018). *Python: La Guía Definitiva para Principiantes para Dominar Python*. Babelcube Inc..
- Yang, M.-S., & Sinaga, K. P. (2019). A Feature-Reduction Multi-View k-Means Clustering Algorithm. *IEEE Access*, vol.7, 114472-114486. doi: 10.1109/ACCESS.2019.2934179.
- Yu, S.-S., Chu, S.-W., Wang, C.-M., Chan, Y.-K., & Chang, T.-C. (2018). Two improved k-means algorithms. *ScienceDirect*, 68(), 747-755. doi:10.1016/j.asoc.2017.08.032.

ANEXOS

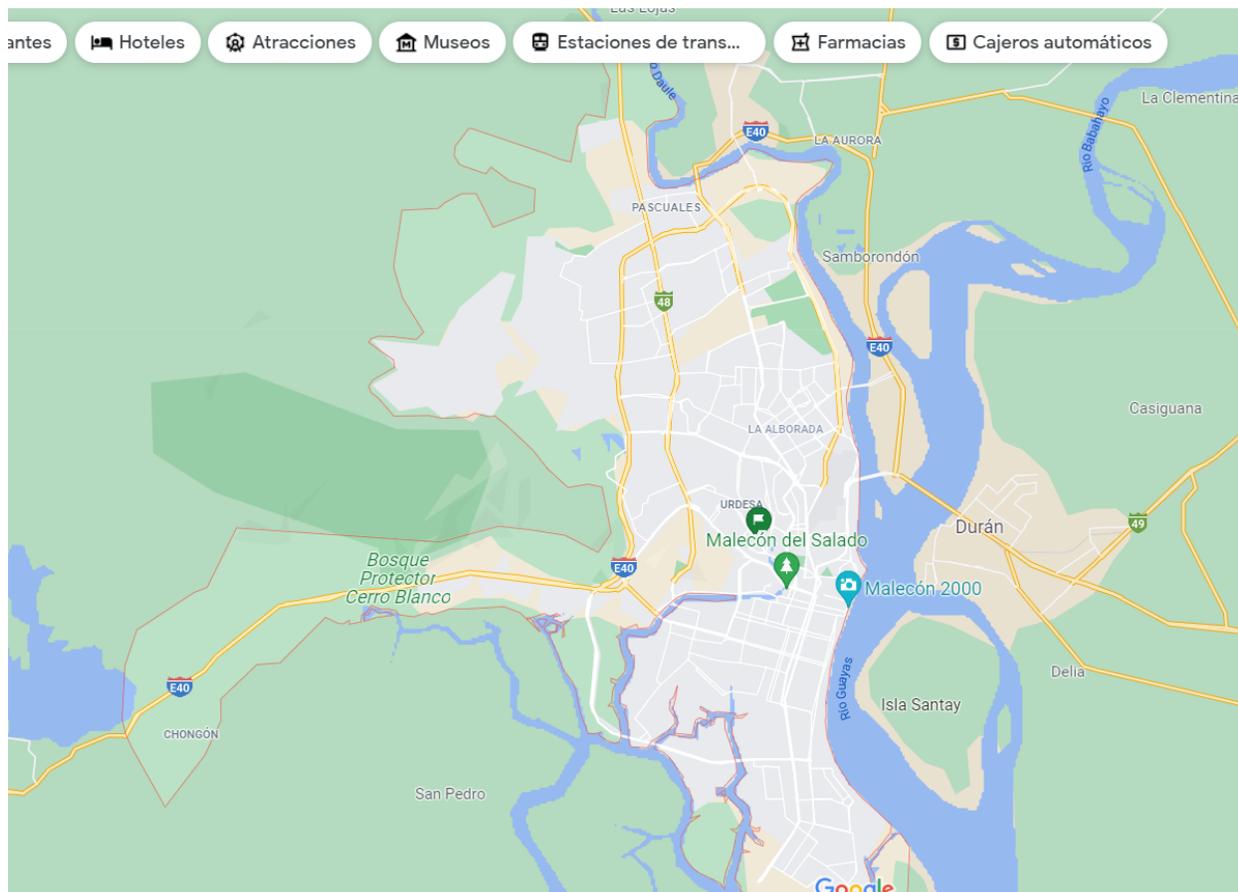
Anexo 1. Planificación de actividades del proyecto

Tareas	Responsable	Estado
SEMANA # 1 (31 MAY - 4 JUN)		
Reunión Tutor: Determinación de requerimientos y objetivos del proyecto.	Ing. Manuel Reyes	Completado
Investigación preliminar	Philco Baque - Serrano Guaman	Completado
Definición preliminar de documentación	Philco Baque - Serrano Guaman	
Desarrollo: Anexo 0, Anexo I y Anexo II	Philco Baque - Serrano Guaman	Completado
SEMANA # 2 (7 JUN - 11 JUN)		
Reunión Tutor: Revisión avance borrador CAP 1.	Ing. Manuel Reyes	Completado
Investigación referenciada al tema propuesto	Philco - Serrano	Completado
Determinación de los requerimientos.	Philco - Serrano	Completado
Entregable: Anexo 0, Anexo I y Anexo II	Philco - Serrano	Completado
SEMANA # 3 (14 JUN - 18 JUN)		
Reunión Tutor: Revisión de cumplimiento de avances	Ing. Manuel Reyes	Completado
Revisión de indicaciones para corrección de CAP 1	Philco - Serrano	Completado
Corrección de CAP 1 y avance de CAP 2	Philco - Serrano	Completado
Investigación de Artículos Relacionados al Proyecto	Philco - Serrano	Completado
SEMANA # 4 (21 JUN - 25 JUN)		
Reunión con Tutor: Revisión de cumplimiento de avances	Ing. Manuel Reyes	Completado
Entregable: Cap 1	Philco - Serrano	Completado
Clasificación de Artículos utilizados en el CAP 2	Philco - Serrano	Completado
Redacción y avance de CAP 2	Philco - Serrano	Completado
SEMANA # 5, # 6 y # 7 (28 JUN - 16 JUL)		
Reunión con Tutor: Revisión de cumplimiento de		

Elaboración: Philco Jorge y Serrano José.

Fuente: Propia.

Anexo 2. Geo-localización del problema



Elaboración: Philco Jorge y Serrano José.
Fuente: Google Maps.

Anexo 3. Fundación Legal

Las Normas Legales en un Proyecto de Titulación

El presente proyecto de titulación se basa en brindar ayuda a los expertos del área logopédica en la toma de decisiones de pacientes con problemas de trastorno dislálico mediante un algoritmo para diagnóstico basado en reglas. Este proyecto fundamenta en la constitución, leyes y normas como se detalla a continuación.

ARTÍCULO DE LA LOES	CONTEXTO
¿Qué regula la LOES? ART. 1 ÁMBITO	Esta Ley regula el sistema de educación superior en el país, a los organismos e instituciones que lo integran; determina derechos, deberes y obligaciones de las personas naturales y jurídicas, y establece las respectivas sanciones por el incumplimiento de las disposiciones contenidas en la Constitución y la presente Ley ARTICULO 1
¿Cuál es el Objeto de esta Ley? ART. 2 OBJETO	Esta Ley tiene como objeto definir sus principios, garantizar el derecho a la educación superior de calidad que propenda a la excelencia, al acceso universal, permanencia, movilidad y egreso sin discriminación alguna.
<u>Entre las funciones</u> ART. 4 DERECHO A LA EDUCACIÓN SUPERIOR	a) Garantizar el derecho a la educación superior mediante la docencia, la investigación y su vinculación con la sociedad, y asegurar crecientes niveles de calidad, excelencia académica y pertinencia; n) Garantizar la producción de pensamiento y conocimiento articulado con el pensamiento universal; y, ñ) Brindar niveles óptimos de calidad en la formación
Principio de Igualdad y Principio de Calidad	El principio de igualdad de oportunidades consiste en garantizar a todos los actores del Sistema de Educación Superior las mismas posibilidades en el acceso, permanencia, movilidad y egreso del sistema, sin discriminación de género, credo, orientación sexual, etnia, cultura, preferencia política, condición socioeconómica o discapacidad. El principio de calidad consiste en la búsqueda constante y sistemática de la excelencia, la pertinencia, producción óptima, transmisión del conocimiento y desarrollo del pensamiento mediante la autocrítica, la crítica externa y el mejoramiento permanente
ART. 87	Como requisito previo a la obtención del título, los y las estudiantes deberán acreditar servicios a la comunidad mediante prácticas o pasantías preprofesionales. debidamente monitoreadas. En los campos de su especialidad, de conformidad con los lineamientos generales definidos por el Consejo de Educación Superior.
ARTÍCULO 19.- DEL REGLAMENTO. - NÓMINA DE GRADUADOS Y NOTIFICACIÓN A LA SENESCYT	Las instituciones de educación superior notificarán obligatoriamente a la SENESCYT la nómina de los graduados y las especificaciones de los títulos que expida, en un plazo no mayor de treinta días contados a partir de la fecha de graduación. (...) este será el único medio oficial a través del cual se verificará el reconocimiento y validez del título en el Ecuador.
ARTÍCULO 144 PRINCIPIOS	Art. 144.- Tesis Digitalizadas. - Todas las instituciones de educación superior estarán obligadas a entregar las tesis que se elaboren para la obtención de títulos académicos de grado y posgrado en formato digital para ser integradas al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

Elaboración: Philco Jorge y Serrano José.

Fuente: Ley Orgánica de Educación Superior.

ARTÍCULO DE LA CONSTITUCIÓN	CONTEXTO
ARTÍCULO 22	Establece: las personas tienen derecho a desarrollar su capacidad creativa, al ejercicio digno y sostenido de las actividades culturales y artísticas, y a beneficiarse de la protección de los derechos morales y patrimoniales que les correspondan por las producciones científicas, literarias o artísticas de su autoría.
ARTÍCULO 26	La educación es un derecho de las personas a lo largo de su vida y un deber ineludible e inexcusable del Estado. Constituye un área prioritaria de la política pública y de la inversión estatal, garantía de la igualdad e inclusión social y condición indispensable para el buen vivir.
ARTÍCULO 28	La educación responderá al interés público y no estará al servicio de intereses individuales y corporativos. Se garantizará el acceso universal, permanencia, movilidad y egreso sin discriminación alguna
ARTÍCULO 350	El sistema de educación superior tiene como finalidad la formación académica y profesional con visión científica y humanista; la investigación científica y tecnológica; la innovación, promoción, desarrollo y difusión de los saberes y las culturas; la construcción de soluciones para los problemas del país, en relación con los objetivos del régimen de desarrollo
ARTÍCULO 355 primer y segundo inciso	El Estado reconocerá a las universidades y escuelas politécnicas autonomía académica, administrativa, financiera y orgánica, acorde con los objetivos del régimen de desarrollo y los principios establecidos en la Constitución

Elaboración: Philco Jorge y Serrano José.

Fuente: Ley Orgánica de Educación Superior.

FACTIBILIDAD LEGAL. – El presente proyecto de titulación es legalmente factible, puesto que no va en contra de ningún lineamiento, normativa o política interna. Se usaron herramientas de código abierto, y aquellas que estén proyectadas en el área de la logopedia sin infringir ningún parámetro legal.

CÓDIGO ORGÁNICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INVENCION	
ARTÍCULO DE LA CONSTITUCIÓN	CONTEXTO
Artículo 104.- Obras susceptibles de protección	La protección reconocida por el presente Título recae sobre todas las obras literarias, artísticas y científicas, que sean originales y que puedan reproducirse o divulgarse por cualquier forma o medio conocido o por conocerse. 12.- SOFTWARE
Artículo 131.- Protección de software	El software se protege como obra literaria. Dicha protección se otorga independientemente de que hayan sido incorporados en un ordenador y cualquiera sea la forma en que estén expresados, ya sea como código fuente; es decir, en forma legible por el ser humano; o como código objeto; es decir, en forma legible por máquina, ya sea sistemas operativos o sistemas aplicativos, incluyendo diagramas de flujo, planos, manuales de uso, y en general, aquellos elementos que conformen la estructura, secuencia y

	<p>organización del programa. Se excluye de esta protección las formas estándar de desarrollo de software. En este sentido, los documentos y textos producidos en las Instituciones de Educación Superior desarrollados con el objeto de obtener sus grados académicos y/o trabajos de facultad, son autores intelectuales con el patrocinio de cada institución, por lo tanto, son acreedores a los derechos de protección intelectual dispuestos en la normativa vigente.</p>
--	---

Elaboración: Philco Jorge y Serrano José.

Fuente: Constitución del Ecuador (2010).

Anexo 4. Criterios éticos a utilizarse en el desarrollo del proyecto

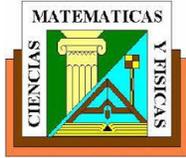
Criterios	Características del criterio	Procedimientos
Credibilidad Valor de la verdad/autenticidad	Aproximación de los resultados de una aproximación frente al fenómeno observado.	<ul style="list-style-type: none"> – Los resultados son reconocidos “verdaderos” por los participantes. – Observación continua y prolongada del fenómeno.
Conformabilidad o Reflexibilidad Neutralidad/Objetividad	Los resultados de la investigación deben garantizar la veracidad de las descripciones realizadas por los participantes.	<ul style="list-style-type: none"> – Transcripciones textuales de las entrevistas. – Contrastación de los resultados de la literatura existente. – Revisión de hallazgos por otros investigadores. – Identificación y descripción de limitaciones y alcances del investigador
Relevancia	Permite evaluar el logro de los objetivos planteados y saber si se obtuvo un mejor conocimiento del fenómeno del estudio.	<ul style="list-style-type: none"> – Configuración de nuevos planteamientos teóricos o conceptuales. – Comprensión amplia del

		<p>fenómeno.</p> <ul style="list-style-type: none"> – Correspondencia entre la justificación y los resultados obtenidos
Adecuación teórica-epistemológica	Correspondencia adecuada del problema por investigar y la teoría existente.	<ul style="list-style-type: none"> – Contrastación de la pregunta con los métodos. – Ajustes de diseño.

Elaboración: Philco Jorge y Serrano José
Fuente: Propia.

Anexo 5. Formatos de técnicas de recolección de datos

Encuesta



UNIVERSIDAD DE GUAYAQUIL

FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS

CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

Proyecto: Algoritmo para diagnóstico de trastorno dislálico en pacientes a partir de la ficha logopédica.

Objetivo: Desarrollar un algoritmo basado en reglas para el diagnóstico de trastorno dislálico en pacientes a partir de las fichas logopédicas, usando herramientas de código abierto con la finalidad de identificar la presencia de los respectivos déficits en los fonemas.

Dislalia

La dislalia es un trastorno del habla donde se presentan dificultades al articular, alteración de determinados sonidos (Dislalia Funcional), sustitución de un sonido por otro de forma incorrecta por deficiencia auditiva (Dislalia Audiógena) y a su vez el trastorno puede ser motivado por causas orgánicas (Dislalia Orgánica). Esta discapacidad es tratada por un Logopeda o terapeuta del lenguaje. De los varios tipos de dislalia solo nos enfocaremos en la funcional para la construcción del algoritmo de diagnóstico.

Dislalia funcional: dificultad para la articulación de uno o varios fonemas (r – s – z – l – k –ch) en niños mayores de 4 años, adolescentes y adultos sin que existan deformaciones anatómicas o fisiológicas en los órganos articulatorios (labio leporino, lengua, encías, maxilares) etc.

1. Antes de haber leído lo anterior. ¿Tenía conocimiento acerca de este tema de la dislalia?

- SI
 - NO
2. ¿Conoce usted a familiares, amigos, infantes o personas cercanas que presenten alteración en alguno de estos fonemas?
- SI
 - NO
3. Conocía usted que los trastornos dislálicos son un problema que causan inseguridad, bullying y problemas de adaptación en infantes, adolescentes y adultos. ¿Cree usted que tendría un impacto positivo adaptar en el área de la logopedia un algoritmo para diagnóstico?
- TOTALMENTE EN DESACUERDO
 - POCO DESACUERDO
 - ACUERDO
 - MUY DE ACUERDO
4. ¿Considera apropiada la construcción de un algoritmo para diagnóstico de trastornos dislálicos?
- SI
 - NO
5. ¿Qué tan importante considera que se genere una herramienta tecnológica para trastornos dislálicos?
- NADA IMPORTANTE
 - POCO IMPORTANTE
 - IMPORTANTE
 - MUY IMPORTANTE
6. ¿Está de acuerdo usted en que se den a conocer estas herramientas tecnológicas en nuestra población?
- TOTALMENTE EN DESACUERDO
 - POCO DESACUERDO

- ACUERDO
 - MUY DE ACUERDO
7. ¿Qué tan relevante cree usted que es dar una solución a estos problemas dislálicos?
- TOTALMENTE IRRELEVANTE
 - POCO IRRELEVANTE
 - RELEVANTE
 - MUY RELEVANTE
8. ¿Se encuentra de acuerdo que al implementar un algoritmo para diagnóstico se llegara a beneficiar tanto al paciente como al especialista Logopeda?
- TOTALMENTE EN DESACUERDO
 - POCO DESACUERDO
 - ACUERDO
 - MUY DE ACUERDO
9. Si usted es una persona afectada. ¿Qué tan seguro se sentiría de usar esta herramienta, para que le genere un diagnóstico del trastorno dislálico?
- TOTALMENTE INSEGURO
 - POCO INSEGURO
 - SEGURO
 - MUY SEGURO

Elaboración: Philco Jorge y Serrano José

Fuente: Propia.

Entrevista



UNIVERSIDAD DE GUAYAQUIL

FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS

CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

Proyecto: Algoritmo para diagnóstico de trastorno dislálico en pacientes a partir de la ficha logopédica.

Nombre del entrevistado: _____

Área de Trabajo: _____

Cargo: _____

PREGUNTAS

1. ¿QUE TEST ES EL QUE MAS UTILIZA PARA DIAGNOSTICAR ESTOS PROBLEMAS DE FONEMAS?

2. SEGÚN SU EXPERIENCIA, ¿CUALES SON LOS FACTORES QUE MAS INFLUYEN EN ESTOS TRASTORNOS DE LENGUAJE?

3. ¿CON QUE FRECUENCIA SE PRESENTAN PACIENTES CON ESTE TIPO DE TRASTORNOS?

4. ¿CUALES SON LOS FONEMAS MAS RECURRENTES QUE SE PRESENTAN EN SUS PACIENTES?

5. ¿QUE TAN PERJUDICIAL PUEDE LLEGAR A SER ESTE TIPO DE TRASTORNOS EN LOS PACIENTES?

6. ¿QUE GRADO DE ASERTIVIDAD BRINDA LA REALIZACIÓN DE LOS TEST EN LOS PACIENTES CON DISLALIA?

7. ¿CUAL ES EL PROCESO A SEGUIR PARA AYUDAR A MEJORAR ESTE TIPO DE TRASTORNOS?

8. ¿CUAL ES EL TIEMPO APROXIMADO PARA PODER TRATAR ESTE TIPO DE TRASTORNOS?

9. ¿SI EXISTIERA UN ALGORITMO PARA DIAGNOSTICO ESTARÍA DISPUESTO A USAR ESTA HERRAMIENTA, PARA QUE LE GENERE UN DIAGNÓSTICO DEL TRASTORNO DISLÁLICO?

Elaboración: Jorge Philco y José Serrano.

Fuente: Propia.

Anexo 6. Validación de expertos.

Juicios de expertos

Para la validación del proyecto se utilizó el instrumento de juicio de expertos con la finalidad de realizar las pruebas de funcionalidad y porcentaje de validación del software desarrollado, adicional los expertos que realicen la validación correspondiente pueda ofrecer valorización para este proyecto y que las técnicas implementadas sean las adecuadas. (Véase Anexo 6: Anexo 6.1, Anexo 6.2 y Anexo 6.3).

CONSTANCIA DE JUICIO DE EXPERTO

Estimado Ingeniero

Manuel Reyes Wagnio

DOCENTE TUTOR DEL TRABAJO DE TITULACIÓN

Ciudad. -

El presente instrumento certifica que se realizó la revisión del proyecto de titulación “ALGORITMO PARA DIAGNOSTICO DE TRASTORNOS DISLALICOS EN PACIENTES A PARTIR DE LAS FICHAS LOGOPEDICAS” cuyos criterios e indicadores empleados permitieron articular el trabajo según se muestra en el Anexo 6.1, por tanto, Philco Baque Jorge Fernando y Serrano Guamán José Miguel estudiante(s) notitulados de la Carrera de Ingeniería en Sistemas computacionales de la Universidad de Guayaquil, pueden continuar con el proceso de titulación en vista que no existen observaciones.

Por lo actuado en el Anexo 6.1, se procede a validar el trabajo de titulación.

Sin otro particular.



Firmado electrónicamente por:
**JOSE LUISALONSO
ANGUIZACA**

Ing. José Luis Alonso Anguizaca, M.Sc.
C.I. N° 0702773052

Elaboración: Philco Jorge y Serrano José.
Fuente: Propia.

CONSTANCIA DE JUICIO DE EXPERTO

Estimado Ingeniero

Manuel Reyes Wagnio

DOCENTE TUTOR DEL TRABAJO DE TITULACIÓN

Ciudad. -

El presente instrumento certifica que se realizó la revisión del proyecto de titulación “ALGORITMO PARA DIAGNOSTICO DE TRASTORNOS DISLALICOS EN PACIENTES A PARTIR DE LAS FICHAS LOGOPEDICAS” cuyos criterios e indicadores empleados permitieron articular el trabajo según se muestra en el Anexo 6.2, por tanto, Philco Baque Jorge Fernando y Serrano Guamán José Miguel estudiante(s) notitulados de la Carrera de Ingeniería en Sistemas computacionales de la Universidad de Guayaquil, pueden continuar con el proceso de titulación en vista que no existen observaciones.

Por lo actuado en el Anexo 6.2, se procede a validar el trabajo de titulación.

Sin otro particular.

**ALFONSO ANIBAL
GUIJARRO
RODRIGUEZ**

Firmado digitalmente por
ALFONSO ANIBAL GUIJARRO
RODRIGUEZ
Fecha: 2021.09.09 07:36:45 -05'00'

Ing. Alfonso Guijarro Rodríguez, Mgs.
C.I. N° 0914312509

Elaboración: Philco Jorge y Serrano José.

Fuente: Propia

CONSTANCIA DE JUICIO DE EXPERTO

Estimado Ingeniero

Manuel Reyes Wagnio

DOCENTE TUTOR DEL TRABAJO DE TITULACIÓN

Ciudad. -

El presente instrumento certifica que se realizó la revisión del proyecto de titulación “ALGORITMO PARA DIAGNOSTICO DE TRASTORNOS DISLALICOS EN PACIENTES A PARTIR DE LAS FICHAS LOGOPEDICAS” cuyos criterios e indicadores empleados permitieron articular el trabajo según se muestra en el Anexo 6.3, por tanto, Philco Baque Jorge Fernando y Serrano Guamán José Miguel estudiante(s) notitulados de la Carrera de Ingeniería en Sistemas computacionales de la Universidad de Guayaquil, pueden continuar con el proceso de titulación en vista que no existen observaciones.

Por lo actuado en el Anexo 6.3, se procede a validar el trabajo de titulación.

Sin otro particular.

**CHARLES
MIGUEL PEREZ
ESPINOZA**  Firmado digitalmente
por CHARLES MIGUEL
PEREZ ESPINOZA
Fecha: 2021.09.11
14:18:16 -05'00'

Ing. Charles Pérez Espinoza MSc.
C.I. N° 0917587743

Elaboración: Philco Jorge y Serrano José.
Fuente: Propia

CONSTANCIA DE JUICIO DE EXPERTO

Estimado Ingeniero

Manuel Reyes Wagnio

DOCENTE TUTOR(A) DEL TRABAJO DE TITULACIÓN

Ciudad. -

El presente instrumento certifica que se realizó la revisión del proyecto de titulación "ALGORITMO PARA DIAGNOSTICO DE TRASTORNO DISLALICO EN PACIENTES A PARTIR DE LA FICHA LOGOPEDICA" cuyos criterios e indicadores empleados permitieron articular el trabajo según se muestra en el Anexo 7, por tanto, Philco Baque Jorge Fernando y Serrano Guamán José Miguel estudiante(s) no titulados de la Carrera de Ingeniería en Sistemas computacionales de la Universidad de Guayaquil, (NO) pueden continuar con el proceso de titulación en vista que (no) existen observaciones.

Por lo actuado en el Anexo 7, se procede a validar el trabajo de titulación.

Sin otro particular.



Lcda. Diana Guerra
C.I. N° 0950025411

Elaboración: Philco Jorge y Serrano José.
Fuente: Propia.

Anexo 7. Tabla del Meta-análisis

N°	Autor(es)	Título	Resumen	Año	Publicación	URL
1	A Fernández Pérez	<i>Orientación a las familias para la prevención de la dislalia funcional</i>	La investigación se realizó con el objetivo de proponer un sistema de actividades de orientación a las familias, para la prevención de la dislalia funcional en los niños de 3-6 años que asisten al Programa “Educa a tu Hijo” del Círculo Infantil: “Los Guerrilleros”, situado en el consejo popular Camacho Libertad, reparto Pedagógico, del municipio de Santa Clara.	2018	<i>dspace.uclv.edu.cu</i>	<i>https://dspace.uclv.edu.cu/handle/123456789/9950</i>

N°	Autor(es)	Título	Resumen	Año	Publicación	URL
2	<i>A Guerrero Sánchez</i>	<i>La dislalia: un estudio de casos en Educación Primaria</i>	La investigación tiene como única finalidad la recuperación del fonema /s/ tras la puesta en marcha de un programa de intervención. Dicha intervención será llevada a cabo con la intención de que el alumno recupere el fonema tanto de forma aislada como en su lenguaje espontáneo. Se realizará de forma que el alumno no se perciba diferente con respecto a sus compañeros, al presentar problemas de comunicación, que no afecte negativamente a su aprendizaje escolar y que, finalmente, no influya en su personalidad.	2017	<i>digibug.ugr.es</i>	https://digibug.ugr.es/bitstream/handle/10481/46271/GUERREROS%C3%81NCH EZ_ARANCHA.pdf?sequence=1

N°	Autor(es)	Título	Resumen	Año	Publicación	URL
3	<i>AJ Lapasaran Pincay, FA Pin Merchán</i>	<i>Estudio de señales de voz mediante electromiógrafos para definir los patrones correctos de la pronunciación y los patrones de pacientes con dislalia.</i>	La dislalia es un trastorno del lenguaje que se presenta a temprana edad y en caso de no ser tratada adecuadamente las secuelas pueden persistir en la edad adulta, los jóvenes que padecen este trastorno suelen ser víctimas de bullying en sus respectivas unidades educativas afectando su estado de ánimo. Este estudio permite a los especialistas en el área terapéutica conocer el comportamiento de los fonemas emitidos por personas que padecen dislalia mediante la construcción de un electromiógrafo basado en Arduino que consiste en un dispositivo conformado por componentes electrónicos cuya utilidad implica ofrecer un diagnóstico más preciso y facilita información de gran utilidad a los especialistas.	2021	<i>repositorio.ug.edu.ec</i>	<i>http://repositorio.ug.edu.ec/handle/redug/52633</i>

N°	Autor(es)	Título	Resumen	Año	Publicación	URL
4	<i>C Cegarra</i>	<i>Marco teórico y proyectos de innovación para alumnos con dislalias</i>	Con este trabajo me he centrado el tema de la dislalia ya que es un tema muy común, debido a que son las alteraciones más conocidas y el trastorno del habla que se presenta con mayor frecuencia en niños de edad escolar. Si escuchamos detenidamente el habla de un niño con trastornos de pronunciación, podemos observar que cambia unos sonidos por otros tota-tola por coca-cola, a veces los omite eses por meses y otras veces quiere imitar el sonido correcto, pero no lo consigue.	2019	<i>lareferencia.info</i>	https://www.lareferencia.info/vufind/Record/ES_8a3a4aa1aaec78eb573cb3d54f87eab6
5	<i>CM Vera</i>	<i>Guía metodológica para la atención a escolares con dislalia funcional en el subnivel básica elemental</i>	El trabajo aborda la necesidad de preparación de los docentes para la atención a escolares con dislalia funcional en el subnivel básica elemental. Tuvo como objetivo, diseñar una guía metodológica que favorezca el aprendizaje de la lectoescritura en los niños y niñas con dislalia funcional.	2020	<i>cienciadigital.org</i>	https://cienciadigital.org/revistacienciadigital2/index.php/CienciaDigital/article/view/1452

Anexo 8. Acta de entrega y recepción

En la ciudad de Guayaquil, septiembre del 2021

Por el presente documento.

Los estudiantes no titulados de la Carrera de Ingeniería en Sistemas Computacionales Philco Baque Jorge Fernando con cédula de identidad N° 0922975149 y Serrano Guamán José Miguel con cédula de identidad N° 0705966935 hacemos la entrega del código fuente del proyecto de titulación a la Dirección de la Carrera de Ingeniería en Sistemas Computacionales en un medio magnético.

Los códigos del programa/producto que se encargaron por compromiso al estar inserto en el proceso de titulación desde fecha 31 de Mayo de 2021.

Para efectos de dar cumplimiento a la entrega del código fuente, cedo todos los derechos de explotación sobre el programa y, en concreto, los de transformación, comunicación pública, distribución y reproducción, de forma exclusiva, con un ámbito territorial nacional.

<hr/>	0922975149
Philco Baque Jorge Fernando	Cédula de identidad N°
<hr/>	0705966935
Serrano Guamán José Miguel	Cédula de identidad N°

Elaboración: Jorge Philco y José Serrano.
Fuente: Propia.

Anexo 10. Evidencias fotográficas

Evidencia de la entrevista aplica a la Lcda. Tairy Pérez

Iniciar Insertar Dibujar Diseñar Disposición Referencias Corrección Revisión Vista Ayuda

UNIVERSIDAD DE GUAYAQUIL
 FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS
 CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

Proyecto: Algoritmo para diagnóstico de trastorno dislálico en pacientes a partir de la ficha logopédica.

Nombre del entrevistado: Lcda. Tairy Pérez
 Área de Trabajo: Lenguaje - Logopedia
 Cargo: Terapeuta del Lenguaje

PREGUNTAS

- ¿QUE TEST ES EL QUE MAS UTILIZA PARA DIAGNOSTICAR ESTOS PROBLEMAS DE FONEMAS?

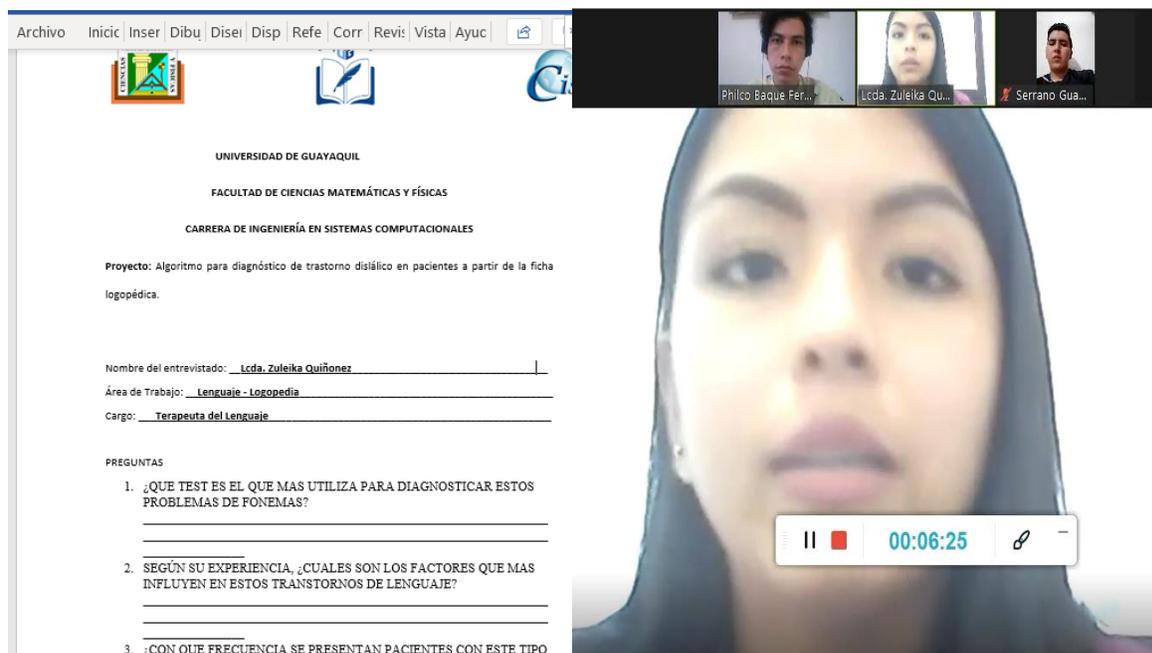
- SEGÚN SU EXPERIENCIA, ¿CUALES SON LOS FACTORES QUE MAS INFLUYEN EN ESTOS TRANSTORNOS DE LENGUAJE?

- ¿CON QUE FRECUENCIA SE PRESENTAN PACIENTES CON ESTE TIPO DE TRANSTORNOS?

Serrano Guamán José
 Jorge Philco Baque
 Tairy Pérez

Nota: En la parte superior izquierda los estudiantes no titulados Serrano José y Philco Jorge seguidos por izquierda inferior la Lcda. Tairy Pérez, en la parte izquierda la Entrevista aplicada a la profesional Logopeda

Evidencia de la entrevista aplica a la Lcda. Zuleika Quiñonez



The image shows a video recording interface. On the left, a document is displayed with the following text:

UNIVERSIDAD DE GUAYAQUIL
FACULTAD DE CIENCIAS MATEMÁTICAS Y FÍSICAS
CARRERA DE INGENIERÍA EN SISTEMAS COMPUTACIONALES

Proyecto: Algoritmo para diagnóstico de trastorno dislálico en pacientes a partir de la ficha logopédica.

Nombre del entrevistado: Lcda. Zuleika Quiñonez
Área de Trabajo: Lenguaje - Logopedia
Cargo: Terapeuta del Lenguaje

PREGUNTAS

1. ¿QUE TEST ES EL QUE MAS UTILIZA PARA DIAGNOSTICAR ESTOS PROBLEMAS DE FONEMAS?

2. SEGÚN SU EXPERIENCIA, ¿CUALES SON LOS FACTORES QUE MAS INFLUYEN EN ESTOS TRANSTORNOS DE LENGUAJE?

3. ¿CON QUE FRECUENCIA SE PRESENTAN PACIENTES CON ESTE TIPO

On the right, a video feed shows three participants in a top bar: Philco Baque Fer..., Lcda. Zuleika Qu..., and Serrano Gua... The main video feed shows a close-up of Lcda. Zuleika Quiñonez. A video control bar at the bottom right of the video feed shows a play button, a red stop button, a timer at 00:06:25, and a share icon.

Nota: En la parte superior izquierda los estudiantes no titulados Serrano José y Philco Jorge seguidos por inferior la Lcda. Zuleika Quiñonez, en la parte izquierda la Entrevista aplicada a la profesional Logopeda

Anexo 9. Artículo científico

ALGORITMO PARA DIAGNÓSTICO DE TRASTORNO DE DISLALIA EN PACIENTES A PARTIR DE LAS FICHAS LOGOPÉDICAS

Reyes-Wagnio Manuel Fabricio¹, Philco-Baque Jorge Fernando²,
Serrano-Guamán José Miguel²
Facultad de Ciencias, Matemáticas y Física, Universidad de
Guayaquil, Ecuador.

¹manuel.reyesw@ug.edu.ec, ²jorge.philcob@ug.edu.ec,
²jose.serranog@ug.edu.ec

Resumen. La Dislalia funcional se define como una alteración del lenguaje, la cual se suele presentar con gran frecuencia en infantes de 5 años de edad en adelante y en la variedad de los casos este problema suele permanecer hasta la edad adulta si no es tratada en su debido momento, las dificultades que se acarrearán con este tipo de trastornos de lenguaje desencadenan severas consecuencias de las cuales podemos mencionar el bullying, discriminación, dificultad de aprendizaje, para poder obtener un diagnóstico de los pacientes que padecen estos trastornos se procedió al desarrollo de un algoritmo para diagnóstico de dislalia funcional, la cual tendrá como objetivo primordial leer las entradas del dataset el cual contendrá las fichas logopédicas realizadas a los pacientes y así poder determinar en qué fonema y en qué posición radica estos problemas de trastornos. Esta necesidad de implementación radica porque en Guayaquil existe muy pocos centros terapéuticos que ayuden a tratar este tipo de trastornos y por su alto costo de consulta lo cual impiden que los niños culminen los tratamientos empezados. Para el desarrollo del algoritmo se realizaron entrevistas a especialistas logopedas y así mismo se recopiló datos de las fichas logopédicas para poderlas adaptar y así tener un diagnóstico correcto, también se utilizó el lenguaje de programación Python 3.9 y según las investigaciones realizadas y los algoritmos seleccionados, el mejor posicionado y de más sencilla aplicación resultó el Basado en Reglas, con él se obtuvo una gran precisión en comparación de los demás algoritmos consultados. Finalmente se muestran las conclusiones donde indican los resultados obtenidos por parte del algoritmo para diagnóstico de dislalia funcional, en donde a partir de estos resultados el especialista logopeda puede tener una noción más clara y así poder decidir cuál sea el tratamiento que se le pueda aplicar al paciente evaluado.

Palabras claves: Fichas logopédicas, Algoritmo de diagnóstico, Basado en reglas, Trastorno dislálico, Dislalia.

1. Introducción

La Dislalia es una alteración fonética la cual consiste en la dificultad de articular, cambiar, omitir o sustituir diferentes fonemas de forma inapropiada por ello es importante desarrollar un prototipo de algoritmo para diagnóstico de dislalia funcional a partir de las fichas logopédicas, donde una vez obtenido los resultados, el algoritmo lo que realizará es brindar un diagnóstico acertado para conocer en cuales fonemas presenta mayor grado de dificultad el paciente, lo cual agilizará el tiempo de respuesta y el porcentaje de precisión en los resultados obtenidos.

¹ Docente de la Carrera de Ingeniería en Sistemas Computacionales de la Universidad de Guayaquil. Ecuador. Facultad de Ciencias, Matemáticas y Física.

² Alumnos de la Carrera de Ingeniería en Sistemas Computacionales de la Universidad de Guayaquil. Ecuador. Facultad de Ciencias, Matemáticas y Física.

1.1. Trabajos Relacionados.

Para los autores (Alarcón, Marchi, Hernández , Chávez, & Hernández Ocaña, 2020) manifiestan en el artículo científico *Clasificadores basados en reglas y selección de atributos para el diagnóstico clínico de subtipos del Síndrome de Guillain-Barré*. Este estudio sirve para analizar el algoritmo basado en reglas, lo cual consiste en utilizar un grupo de normas IF-THEN para la categorización. Una regla IF-THEN es una expresión de la manera IF condition THEN conclusión. La parte IF de una regla se sabe cómo el precedente o precondition de la regla. La parte THEN es la regla consecuente. Estos tienen algunas ventajas, es enormemente expresiva, simple de interpretar y de producir; puede clasificar novedosas instancias y el rendimiento se compara con árboles de decisiones. También utilizaron algoritmos de aprendizaje automático como por ej.: JRip, OneR y PART, en lo cual obtuvieron resultados a través de los promedios que arrojaron estos algoritmos, los cuales fueron entre 0.4923 a 0.6265, siendo cómo mayor porcentaje el algoritmo JRip logrando balancear el resultado y obteniendo un porcentaje de confiabilidad reformado.

En el trabajo del autor (Espinoza, 2018) indica en su tesis de grado *Diseño y desarrollo de una herramienta lúdica informática para estimular e intervenir los niveles de lenguaje en niños desde los 4 a 5 años*. La tesis consta de una aplicación multiplataforma, la cual posibilita excitar los diferentes niveles de lenguaje según las necesidades de cada infante, para poder hacer esto se desarrolló un sistema recomendador el cual posibilita examinar los datos del infante y sugerir tal cual la terapia óptima que se debería hacer dentro del programa de estimulación. Se consultaron expertos en la estimulación infantil, especialistas y psicólogos infantiles, para poder conocer bajo qué parámetros se debía desarrollar el sistema. Las herramientas que se utilizaron fueron lenguaje Python, minería y análisis de datos aportados por la librería Scikit-learn.

Los autores (Regalado Macias & Calle De la A, Construcción de algoritmo de similitud para el sistema del Proyecto FCI, Herramienta para el tratamiento de trastornos fonéticos (Dislalia), 2020) resaltan en su tesis de grado *Construcción de algoritmo de similitud para el sistema del proyecto FCI herramienta para el tratamiento de trastornos fonéticos (Dislalia)*. En esta propuesta utilizaron una metodología de prueba experimental para el desarrollo del algoritmo, a su vez mediante entrevistas y encuestas a personas afines al estudio determinaron que el conocimiento en informática es bajo, y que desconocían que la tecnología puede ser usada en beneficio médico, y a su vez el algoritmo de similitud desarrollado ayuda a los especialistas a obtener una información óptima y precisa para pacientes con Dislalia.

2. Metodología

El trabajo realizado, tiene como principal tipo de investigación a la investigación descriptiva mediante un diseño de corte transversal. Se usó también una metodología Kanban.

La investigación descriptiva la cual Burgo, León, Cáceres, Pérez y Espinoza (Burgo, León, Cáceres, Pérez, & Espinoza, 2019) mencionan que son aquellas que observan y describen el fenómeno; no requieren la manipulación de variables. Se desarrollan para conocer las características más relevantes del estado del fenómeno y los factores asociados al problema. Cabe mencionar (Galeano, 2020) que la presente investigación se ejecuta con una modalidad de campo bibliográfica. En la indagación de campo se ha

realizado una entrevista para conocer la necesidad de implementar un algoritmo para diagnóstico de trastorno dislálico. Por otro lado, del total de la indagación, se hizo alrededor de un 70% de revisiones bibliográficas de fuentes electrónicas, con el propósito de sustraer información importante que sustente la ejecución de la investigación y el proyecto, el 30% de la modalidad es de campo, debido a que se desarrollara un prototipo de algoritmo basado en reglas que consiste en realizar una condición y presentar los procesos sucesivamente hasta terminar con los diferentes diagnósticos dados para cada paciente asignado.

Según lo indica Gilberts (Gilberts, 2020), el concepto Kanban ha pasado a conformar parte de las denominadas metodologías ágiles, cuyo objetivo es gestionar de forma general cómo se van completando las labores. Kanban es un vocablo japonés que significa “tarjetas visuales”, donde Kan es “visual”, y Ban corresponde a “tarjeta”. Esta metodología se la utiliza para poder tener una mejor organización al momento de enlistar las actividades a realizar durante el proyecto y así poder terminar en el tiempo indicado cada actividad y evitar los cuellos de botellas. La metodología Kanban se divide en tres fases, las cuales son:

- **Request.** - Planificación de las actividades a realizar al momento de desarrollar el sistema.
- **In Progress.** - Las actividades que están en progreso durante las etapas del proyecto.
- **Done.** - Las actividades que ya se han realizado durante la realización del proyecto.

El diseño de corte transversal se cataloga como un análisis observacional de base personal que frecuenta tener un doble objetivo: detallado y analítico. Además, es conocido como análisis de prevalencia o encuesta transversal; su objetivo fundamental es detectar la frecuencia de una condición o patología en la población estudiada y pertenece a los diseños básicos en epidemiología al igual que el diseño de casos y controles y el de cohortes (estudio al que se observa a lo largo de un lapso) según lo indican Rodríguez y Mendivelso (Rodríguez & Mendivelso, 2018). Se utilizó este diseño porque se analizará dos variables, que son cuantitativas y cualitativas:

Cuantitativas: Son los fonemas que forman el dataset y están determinados en el rango 0 - 1.1 - 1.2 - 1.3 - 1.4 - 2.1 - 2.2 - 2.3 - 3.1 - 3.2 - 3.3 - 4.1 - 4.2 - 4.3 - 5.1 - 5.2 - 5.3

Cualitativas: Al momento de ingresar al diagnóstico el dataset se convierte en variables no medibles para poder trabajar con dicha información.

3. Materiales y Métodos.

3.1. Inteligencia artificial.

Machine Learning

El cual es un software de búsqueda y análisis de artículos científicos. Los parámetros de búsqueda se efectuaron utilizando

Naive Bayes

Uno de los algoritmos a tomar en cuenta para el desarrollo del proyecto es Naive Bayes, el cual se encuentra dentro del grupo de algoritmos de aprendizaje supervisado. Este asume que cada una de las características que utiliza son condicionalmente independientes entre sí dada alguna clase. Es útil para conjuntos de datos muy grandes y se puede utilizar el algoritmo de clasificación Naive Bayes con un conjunto de datos pequeño, pero la precisión y la recuperación serán muy bajas. Dado a que se fundamentan en la libertad de las propiedades con dichos algoritmos se pierde la

función de explotar las interrelaciones entre estas según indica Shaier (2019).

Algoritmo Basado en Reglas

Las reglas forman parte del dialecto humano, lo que hace que estas reglas resulten fáciles para la comprensión y entendimiento del conocimiento cotidiano, lo cual hace que para este apartado resulte de bastante relevancia puesto que, al ser más interactivo con el habla humana, se podrá utilizar mejor la clasificación para así resolver la problemática establecida.

3.2 Herramientas tecnológicas.

Publish or Perish

Para la recolección de datos se investigó mediante la herramienta “Publish or Perish”, la cual es un software de búsqueda y análisis de artículos científicos. Los parámetros de búsqueda se efectuaron utilizando las palabras claves establecidas en el proyecto, encontrando diversos artículos que cumplan con las coincidencias de búsqueda establecidas e identificando diversas fuentes bibliográficas las cuales se fueron analizando, para realizar una hoja de cálculo que se procedió a examinar posteriormente, la cual está constituida por autores, título, año, publicación y artículo URL. Según (Galvez Quilca, SISTEMA WEB PARA EL PROCESO DE CONTROL DE PROYECTOS EN LA EMPRESA HUNDRED S.A.C., 2019) indica que la implementación de las técnicas de recolección de datos contempla elaborar y realizar un procedimiento bien ordenado que posibilite recopilar estos datos con un objetivo definido. Así mismo la encuesta según (Arenal Laza, 2019) es el instrumento más usado y conocido una vez que se desea conseguir precisión y representatividad partiendo de manera directa de las consideraciones personales y no estructurales, para lograr una conclusión sobre los sujetos insertos en ellos.

Python 3.9

Python es un lenguaje de programación cuya ventaja es poder ser usado en cualquier plataforma al mismo tiempo que el código es intuitivo. Posee muchas versiones y es capaz de interactuar con otros lenguajes de programación de tal forma que puede ejecutar casi cualquier programa que se desee, define Walker (Walker, 2018). Algunas de las características de Python que se pueden mencionar, es que es un lenguaje multiplataforma, tanto para Mac, Windows y Linux y es muy ventajoso para ejecutar su código fuente, por lo que se adapta sin muchos cambios, su lenguaje es interpretado, usa tipado dinámico, fuertemente tipado. Soporta varios paradigmas de programación como, estructurada, orientada a objetos, y en menor medida programación funcional. La indentación (formato del código) en Python se da forma estructural. Este fue el lenguaje elegido para la elaboración de los algoritmos elegidos.

STAT::FIT

Para Carrasco, Ramos y Morales (Carrasco-Machado, Ramos, & Morales-Muñoz, 2020) , define a Stat-fit como un software que, desde las hojas de cálculo, ingreso manual o por archivos de texto; toma datos brutos para convertir estos en la distribución de probabilidad óptima para el programa ProModel. Mediante esta herramienta se obtuvo aquellos datos estadísticos correspondientes a la simulación realizada para la obtención de más data que ayude en el entrenamiento de los algoritmos.

Excel

Excel es una hoja de cálculo que se centra en el trabajo de tipo numérico. Según lo

define Antúñez (Antúñez, 2017) ,se encuentra dentro de la Suite Microsoft Office, es decir, está dentro de un conjunto de aplicaciones dirigidas a la Ofimática, esto sin la necesidad de tener un ordenador con grandes capacidades puesto que los requisitos para este tipo de programas no son elevados.

4. Caso de estudio.

Mediante la data obtenida, se aplicaron diversos algoritmos con la finalidad de obtener un diagnóstico que garantice confiabilidad y detección de aquellos posibles pacientes de dislalia. Para ello, se contó con la metodología propuesta previamente en la cual se desarrolló en fases.

Fase 1: Diseño Base de Datos - Dataset

Se pudo obtener acceso a una base de datos dentro del proyecto FCI de la Facultad de Ciencias Matemáticas y Físicas, levantada con 50 datos reales de pacientes con diferentes tipos de dislalias y diferentes atributos, para la construcción del algoritmo para diagnóstico este proyecto se ha enfocado en la dislalia funcional por lo cual para la selección de los atributos/variables más significativos de la base de datos realizamos lo siguiente:

4. Investigaciones dentro de artículos científicos, revistas científicas, libros sobre dislalia, fichas logopédicas que comprenden las pruebas de evaluación ELA-R, PAF, entrevista a un especialista logopeda el cual con su criterio ayudó a descartar y seleccionar las variables más optimas en la toma de decisión del diagnóstico de la dislalia funcional.
5. Lo siguiente fue crear un dataset en Excel a partir de la base de datos, el cual está conformado por:

Variables: [M -P- B- D- T- Z- F- S- N- L- R- RR- LL- Ñ- K- G- J- X- BL- CL- FL- GL- PL- BR- CR- DR- PR- GR]

Datos para cada variable del déficit de fonemas dentro de dislalia funcional:

Sustitución inicio, medio, final [1.1 – 1.2 – 1.3]

Omisión inicio, medio, final [2.1 – 2.2 - 2.3]

Distorsión inicio, medio, final [3.1 – 3.2 – 3.3]

Inversión inicio, medio, final [4.1 – 4.2 – 4.3]

Adición inicio, medio, final [5.1 - 5.2 – 5.3]

y los datos entorno a las pruebas realizados a los pacientes que ayudarían al proceso del diagnóstico de la dislalia funcional.

6. Por último, contando con los datos y variables óptimas en este dataset realizado en Excel, se almaceno un archivo con extensión .csv para su posterior uso en el algoritmo construido en Python.

Fase 2: Simulación de Datos

Una vez selecciona las variables y creado el nuevo dataset, se resolvió que los datos en la base real de pacientes con dislalia funcional eran muy pocos, para lo cual se realizó una simulación de 1002 datos, tomando los datos reales y por consiguiente aumentando más casos para el dataset.

Las herramientas utilizadas para la simulación de estos datos fueron dos: Microsoft Excel el cual tiene integrado Microsoft Visual Basic para aplicaciones, herramienta mediante la cual se generarán los 1002 datos de forma aleatoria. Para llevar a cabo esta simulación se utilizará el método Monte Carlo ayudados de la herramienta @Risk 8.1 que es un complemento de Excel usado para la aplicación de simulación.

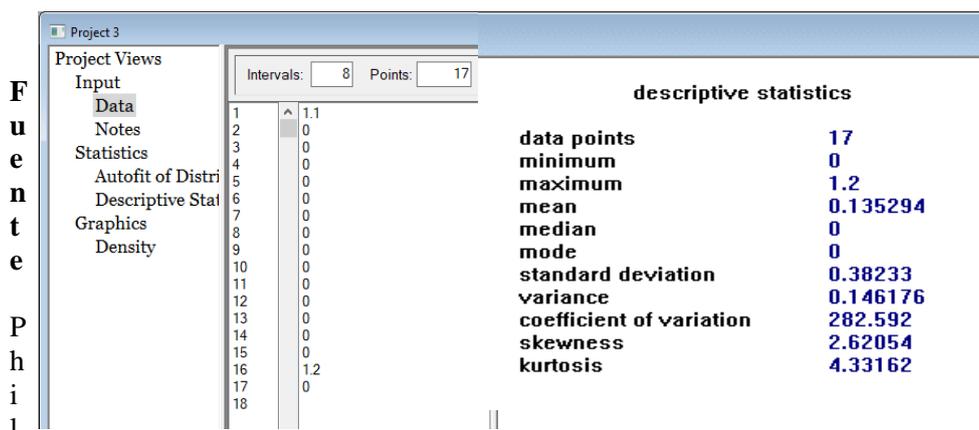
Simulación Monte Carlo

El propósito primordial de la simulación de Montecarlo es intentar emular la conducta de variables reales para, en el tamaño de lo viable, examinar o pronosticar cómo van a

cambiar. La simulación Monte Carlo es un instrumento estadístico que posibilita la modelación de resultados conforme con la conducta histórico de los datos y su posibilidad de ocurrencia, (Salazar & Alzate, Aplicación de la simulación Monte Carlo en la proyección del estado de resultados. Un estudio de caso, 2018)

Cabe aclarar que para poder utilizar el método Monte Carlo se debe conocer las distribuciones de cada una de las variables del dataset, esto lo logramos haciendo uso de la herramienta STAT::FIT, esta herramienta tecnológica permite la entrada de mínimo 10 y máximo 50 datos, dado el caso que el dataset contiene los datos reales de pacientes estos son ingresados en la herramienta por variable.

Figura 1. STAT::FIT. Distribución de la variable M



co, J, Serrano, J

Elección del algoritmo

Se midieron diversos algoritmos de machine learning en este caso el algoritmo Naive bayes, al cual mediante su entrenamiento arroja un resultado de 0.13% de precisión como indica la figura 2, árbol de decisión con un porcentaje de 0.35 % como indica la figura 3 y K-NN con un porcentaje de 0.27% dado el motivo que el dataset construido contaba con muchas entradas y esto daba muchas combinaciones de salida, y no existía similitud para que el algoritmo pueda ser entrenado y realice la debida predicción del diagnóstico. Al no obtener un resultado positivo en la exactitud mediante algoritmos de machine learning, se realizó una investigación y se eligió algoritmos basados en reglas el cual va a ser utilizado en el desarrollo de este prototipo ya que con ello se pudo satisfacer el objetivo principal que es generar un diagnóstico a partir del dataset. El algoritmo basado en reglas incluyó un porcentaje de precisión de 86%.

Figura 2. Algoritmo Naive Bayes

```
[45] #Precisión del modelo Naive Bayes
p2=accuracy_score(y_test, y_pred2)
print('Precisión del modelo')
print()
print(p2)

Precisión del modelo
0.13545816733067728
```

Fuente Philco, J, Serrano, J

Figura 3. Algoritmo Árbol de decisión.

```
[33] # exactitud del modelo
      from sklearn.metrics import accuracy_score
      accuracy_score(y_test, y_fit)
```

```
0.35323383084577115
```

Fuente Philco, J, Serrano, J

Figura 4. Algoritmo K-NN

```
[40] #Se entrena el modelo
      algoritmo.fit(x_train, y_train)
      #Se obtienen los resultados del entrenamiento
      y_pred = algoritmo.predict(x_test)

[41] from sklearn.metrics import accuracy_score
      accuracy_score(y_test, y_pred)

0.2749003984063745
```

Fuente Philco, J, Serrano, J

Fase 3: Diseño del algoritmo

- **Base de datos:** se trabajó con 1002 datos alojados en el dataset (.csv)
- **Preparación ambiente de trabajo:** Se instalaron herramientas como el editor de código Visual Studio Code que permitió con su enorme cantidad de extensiones gratuitas poder desarrollar en el lenguaje Python, de forma amigable el algoritmo para diagnóstico.

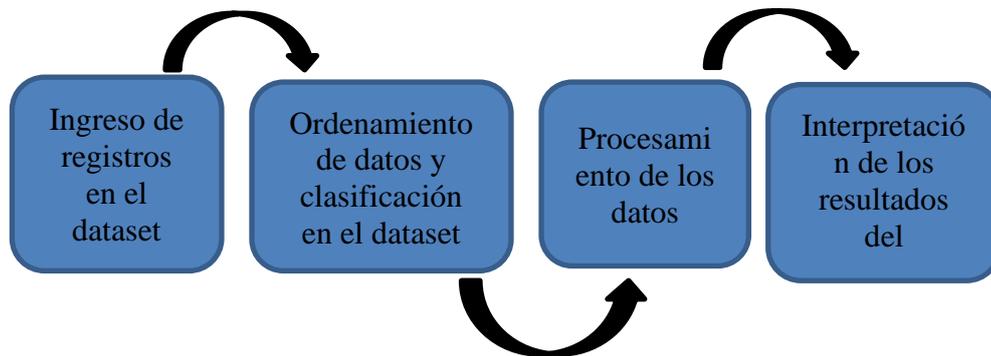
Figura 5. Fragmentos de código para construcción del algoritmo

```
1 import tkinter as tk
2 import pandas as pd
3 import time
4 data = pd.read_csv('C:\\Users\\hp240\\Desktop\\juk\\PRUEBA.csv', sep=';', encoding='latin-1')
5 datos = data.values
6 columna = data.columns
7 columna=columna.delete(0)
8 columna=columna.delete(0)
9 columna = columna.values
10
11
12
13 class Example(tk.Frame):
14     def __init__(self, parent):
15         tk.Frame.__init__(self, parent)
16         self.prompt = tk.Label(self, text="Las fichas logopedicas cargadas al algoritmo son")
17         self.entry = tk.Entry(self)
18         self.output = tk.Label(self, text="")
19         self.submit = tk.Button(self, text="Diagnosticar", command = self.diagnosticar)
20         self.prompt.pack(side="top", fill="x")
21         self.entry.pack(side="top", fill="x", padx=20)
22         self.submit.pack(side="bottom")
23
24
```

Fuente Philco, J, Serrano, J

Implementación

En los alcances del proyecto la implementación del algoritmo consiste en la construcción del algoritmo en Python, y el prototipo donde se introduzcan las entradas de los resultados más destacados de las fichas logopédicas y arroje un diagnóstico.

Figura 6. Proceso de implementación del algoritmo

Fuente Philco, J, Serrano, J

Pruebas Inducidas

En este apartado se medirá la efectividad del resultado del diagnóstico que otorgue el algoritmo, mediante pruebas inducidas en la cual se cambiarán los valores ingresados en el dataset y corroborar que el resultado sea el que ya tendremos conocimiento y así poder verificar que todas las reglas que se encuentran inmersas están correctamente formuladas dentro de este prototipo.

Prueba N°1

Figura 7. Dataset con datos de pacientes a diagnosticar

A	B	C	D	E	F	G	H
Paciente	edad	M	P	B	D	T	Z
1	12	0	0	0	0	0	0
2	6	0	1.2	1.2	2.3	4.3	2.1
3	8	2.1	4.1	3.2	0	4.1	5.2
4	9	5.2	3.3	0	3.2	1.2	1.1
5	7	3.3	5.2	2.2	2.2	0	0

Philco, J, Serrano, J

Primero se probará el algoritmo ingresando un Dataset con 5 pacientes como indica la Figura 7, y los fonemas a evaluar serán solo 6, y se pedirá que arroje el diagnóstico del paciente 1, en el cual como resultado arrojará que no presenta ningún inconveniente en ninguna de los fonemas existentes como indica la figura 8.

Figura 8. Salida por consola del diagnóstico paciente 1

```

Las fichas logopedicas cargadas al algoritmo son de 5 pacientes
Ingrese el numero del paciente a diagnosticar dentro del rango: 1
Datos cargados del test de evaluacion
Paciente 1
Los resultados del diagnostico son:
Sustitucion: ninguna,
Omision: ninguna,
Distorsion: ninguna,
Inversion: ninguna,
Adicion: ninguna
Digite 1 para otro diagnostico, presione cualquier otra tecla para finalizar: : █
  
```

Fuente Philco, J, Serrano, J

Como se ha mostrado previamente, el diagnóstico en el paciente 1 es el acertado, no presenta ningún tipo de trastornos en ninguna de los fonemas, se lo puede corroborar en la imagen, que indica cuales son la clasificación de los trastornos, pero en ninguno de ellos el paciente evaluado presenta inconvenientes. Así mismo se va a diagnosticar con el mismo dataset al paciente número 3 en el cual ya sabremos que deben presentar varios tipos de trastornos en el diagnóstico final.

Figura 9. Salida por consola del diagnóstico paciente 3

```

Las fichas logopedicas cargadas al algoritmo son de 5 pacientes
Ingrese el numero del paciente a diagnosticar dentro del rango: 3
Datos cargados del test de evaluacion
Paciente 3
Los resultados del diagnostico son:
Sustitucion: ninguna,
Omision['Fonema: M', 'Posicion: inicio'],
Distorsion['Fonema: B', 'Posicion: medio'],
Inversion['Fonema: P', 'Posicion: inicio', 'Fonema: T', 'Posicion: inicio'],
Adicion['Fonema: Z', 'Posicion: medio']

```

Fuente Philco, J, Serrano, J

De la misma forma se puede visualizar como indica la Figura 9 los resultados obtenidos son los esperados en cada uno de los fonemas a evaluar, es decir, diagnostica que tipo de déficit de fonema presenta y en qué posición se encuentra la dificultad, en el cual se podría acotar que el algoritmo cumple con las expectativas propuestas.

Prueba N°2

Figura 10. Dataset con datos de pacientes a diagnosticar

Paciente	edad	M	P	B	D	T	Z	F	S	N	L	R	RR
1	12	6.1	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3
2	6	0	0	0	0	0	0	0	0	0	0	0	0
3	8	2.1	4.1	3.2	0	4.1	5.2	0	1.1	2.1	0	0	3.1
4	9	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1
5	7	3.3	5.2	2.2	2.2	0	0	1.3	3.3	2.2	4.1	0	5.3

Fuente Philco, J, Serrano, J

En esta segunda prueba se le ingresara al algoritmo un mismo dataset con 5 pacientes, y los fonemas (variables de entrada) a evaluar ya se incrementaron a 12, y en este caso se pedirá que arroje el diagnóstico del paciente 2, en el que obtendremos como resultado que no presenta ningún inconveniente en ninguna de los fonemas evaluados.

Figura 11. Salida por consola del diagnóstico paciente 2

```

Las fichas logopedicas cargadas al algoritmo son de 5 pacientes
Ingrese el numero del paciente a diagnosticar dentro del rango: 2
Datos cargados del test de evaluacion
Paciente 2
Los resultados del diagnostico son:
Sustitucion: ninguna,
Omision: ninguna,
Distorsion: ninguna,
Inversion: ninguna,
Adicion: ninguna

```

Fuente Philco, J, Serrano, J

Se visualiza en la Figura 11, que los resultados obtenidos en este paciente no presentan

ninguna dificultad en los fonemas evaluados, lo cual se puede afirmar decir que al aumentar la cantidad de fonemas y darle una categorización el algoritmo sigue brindando un diagnóstico correcto. De la misma manera con el mismo dataset ingresado también se evaluará al paciente número 4 en el cual como resultado dará que presentará Omisión en todos los fonemas en la posición inicial.

Figura 23. Salida por consola del diagnóstico paciente 4

Fuente
Philco,
J.,
Serrano,
J.

```
Ingrese el numero del paciente a diagnosticar dentro del rango: 4
Datos cargados del test de evaluacion
Paciente 4
Los resultados del diagnostico son:
Sustitucion: ninguna,
Omisión['Fonema: M', 'Posición: inicio', 'Fonema: P', 'Posición: inicio', 'Fonema: B', 'Posición: inicio', 'Fonema: D', 'Posición: inicio', 'Fonema: T', 'Posición: inicio', 'Fonema: Z', 'Posición: inicio', 'Fonema: F', 'Posición: inicio', 'Fonema: S', 'Posición: inicio', 'Fonema: N', 'Posición: inicio', 'Fonema: L', 'Posición: inicio', 'Fonema: R', 'Posición: inicio', 'Fonema: RR', 'Posición: inicio'],
Distorsion: ninguna,
Inversion: ninguna,
Adición: ninguna
```

Como se muestra en la Figura 23, el resultado de diagnóstico fue el esperado, el paciente 4 refleja que presenta déficit de fonemas por Omisión, en todos los fonemas en la posición inicial, dando una vez más comprobado que el algoritmo basado en reglas nos brinda un óptimo diagnóstico.

Prueba N°. 3

Figura 24. Dataset con datos de pacientes a diagnosticar

Fuente
Philco,
J.,
Serrano,
J.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD			
1	Paciente	edad	M	P	B	D	T	Z	F	S	N	L	RR	LL	K	G	J	X	BL	CL	FL	GL	PL	BR	CR	DR	RR	GR					
2	1	2	61	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43
3	2	6	0	12	12	23	43	21	13	31	0	0	11	22	0	12	52	42	11	32	33	0	22	41	0	12	12	0	0	0	0	0	
4	3	8	21	41	32	0	41	52	0	11	21	0	0	31	53	33	0	0	12	23	32	0	41	51	0	22	22	0	0	22			
5	4	9	52	33	0	32	12	11	0	12	0	0	0	0	22	21	31	33	0	0	12	42	53	0	0	0	22	32	0	42			
6	5	7	11	11	11	11	11	11	11	11	11	11	11	11	11	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Finalmente, en esta tercera prueba se le ingresara al algoritmo un mismo dataset con 5 pacientes, y los fonemas a evaluar ya serán los utilizados en las fichas logopédicas siendo estos una cantidad de 28, y en este caso se pedirá que arroje el diagnóstico del paciente 5, en el que obtendremos como resultado que dicho paciente presentará trastornos de dislalia en los primeros 14 fonemas, en donde se clasificara como una sustitución en la posición inicial de estos.

Figura 25. Salida por consola del diagnóstico del paciente 5

Fuente
Philco,
J.,
Serrano,
J.

```
Las fichas logopedicas cargadas al algoritmo son de 5 pacientes
Ingrese el numero del paciente a diagnosticar dentro del rango: 5
Datos cargados del test de evaluacion
Paciente 5
Los resultados del diagnostico son:
Sustitucion['Fonema: M', 'Posición: inicio', 'Fonema: P', 'Posición: inicio', 'Fonema: B', 'Posición: inicio', 'Fonema: D', 'Posición: inicio', 'Fonema: T', 'Posición: inicio', 'Fonema: Z', 'Posición: inicio', 'Fonema: F', 'Posición: inicio', 'Fonema: S', 'Posición: inicio', 'Fonema: N', 'Posición: inicio', 'Fonema: L', 'Posición: inicio', 'Fonema: R', 'Posición: inicio', 'Fonema: RR', 'Posición: inicio', 'Fonema: LL', 'Posición: inicio', 'Fonema: \x04', 'Posición: inicio'],
Omisión: ninguna,
Distorsion: ninguna,
Inversion: ninguna,
Adición: ninguna
```

El resultado obtenido sigue siendo el acertado, se evaluaron los 28 fonemas al paciente 5 y solo presenta sustitución en los primeros 14 fonemas y todos en la posición inicial, tal como se lo indico anteriormente, dejando una vez más aclarado que el algoritmo brinda un diagnóstico certero en todas las pruebas realizadas.

5. Resultados.

El algoritmo basado en reglas y el de aprendizaje supervisado Naive Bayes fueron empleado con el uso de la herramienta tecnológica Python, como es de conocimiento general este lenguaje de programación es uno de los más utilizados a nivel mundial y se adapta en gran parte a los requerimientos que se lo aplique. La codificación de estos algoritmos fue desarrollada bajo el entorno Visual Studio Code, cuya finalidad es la obtención del diagnóstico de trastornos dislálicos a partir de la información obtenida del dataset.

Como primer resultado se obtuvo que, a través de la aplicación del algoritmo Naive Bayes no se otorgaba un diagnóstico esperado, por tal motivo no fue recomendable su uso. Por otra parte, con el algoritmo basado en reglas se obtuvo un éxito bastante favorable en comparación con el algoritmo anterior, dándole así mayor relevancia y siendo finalmente este último el más indicado para su desarrollo y aplicación para un diagnóstico aceptable, agregando que, los tiempos de procesamiento también son bastante cortos lo cual lo hace que sea un algoritmo eficiente.

1. Mediante investigaciones científicas consultadas se pudo obtener toda la información más relevante que aportaron a definir los tipos de trastornos dislálicos y conocer cuáles son sus principales causas, tras ello, se usaron estos conocimientos para poder estructurar de manera adecuada el algoritmo basado en reglas.
2. Para el dataset, se obtuvo una base de datos brindada por fuentes primarias, que sirvió como guía para posteriormente también poder realizar una simulación y depuración de datos haciendo uso de la herramienta tecnológica @risk 8.3.
3. Finalmente, habiendo estructurado el dataset, se prosiguió con la construcción del algoritmo basado en reglas, en el cual se obtuvieron los resultados esperados, cumpliendo con las expectativas propuestas y dando por afirmado que el diagnóstico resultante es correcto.

6. Conclusiones.

1. Según la comparativa de los diferentes tipos de algoritmos investigados como lo son: basado en reglas, aprendizaje supervisado como Naive Bayes, K-NN, arboles de decisiones, se pudo analizar mediante indicadores cuál de estos algoritmos cuenta con mayor precisión, arrojando como resultado que el más óptimo al momento de generar el diagnóstico es el basado en reglas en relación con la base de datos utilizada.
2. Se realizó una simulación de los datos utilizando herramientas tecnológicas tales como @Risk 8.3, Stat::Fit para ser utilizados dentro del presente proyecto, a su vez se escribió un algoritmo el cual será el encargado de generar un diagnóstico de trastornos dislálicos en cada uno de los pacientes a evaluarse.
3. Finalmente, se redactó el artículo científico que contiene todas las etapas de la investigación de este trabajo dejando así una memoria informativa, con el cual se pretende ayudar a futuros trabajos pues se menciona el desarrollo del algoritmo basado en reglas y los casos de trastornos fonológicos que comprende la dislalia funcional.

7. Referencias.

- D'Angiolo, F. G., Kwist, I. F., Loiseau, M., Contreras, D. E., & Asteasuain, F. (Octubre de 2019). *Repositorio Institucional de la UNLP*. Obtenido de Algoritmos de regresión lineal aplicados al mantenimiento de un datacenter: <http://sedici.unlp.edu.ar/handle/10915/91401>
- Alarcón, D., Marchi, J., Hernández, J., Chávez, O., & Hernández Ocaña, B. (2020). Clasificadores basados en reglas y selección de atributos para el diagnóstico clínico de subtipos del Síndrome de Guillain-Barré. *rsc.cic.ipn.mx*, 1-13. Obtenido de https://rsc.cic.ipn.mx/2020_149_8/Clasificadores%20basados%20en%20reglas%20y%20seleccion%20de%20atributos%20para%20el%20diagnostico%20clinico.pdf
- ALBOR-COHS, G. (1999). *ELA-R: Examen logopedico de articulación (revisado)*. Obtenido de ACADEMIA: https://www.academia.edu/8875885/ELA_Manual_de_Referencia_ALBOR_COHS
- Ali, M., Jung, L. T., Abdel-Aty, A.-H., Abubakar, M. Y., Elhoseny, M., & Ali, I. (2020). Semantic-k-NN algorithm: An enhanced version of traditional k-NN algorithm. *Expert Systems with Applications vol.151*, 9, doi: <https://doi.org/10.1016/j.eswa.2020.113374>.
- Antúñez, C. (2017). *Iniciación a excel 2013*. Editorial Elearning, SL.
- aprendemachinelearning. (10 de Julio de 2018). *Clasificar con K-Nearest-Neighbor ejemplo en Python*. Obtenido de KNN: <https://www.aprendemachinelearning.com/clasificar-con-k-nearest-neighbor-ejemplo-en-python/>
- Ardila, E. (2020). COVID-19 en Colombia e inmunidad de rebaño: ¿es momento de considerarla? *Revista Colombiana de Endocrinología, Diabetes y Metabolismo*, 57-59.
- Arenal Laza, C. (2019). *Investigación y recogida de información de mercados*. Tutor Formación. Obtenido de https://books.google.com.ec/books?id=srenDwAAQBAJ&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false.
- Auquilla Morocho, E. F. (2021). *MINERÍA DE OPINIÓN PARA TEXTOS EN ESPAÑOL USANDO PROCESAMIENTO NATURAL DEL LENGUAJE*. Quito: Universidad Politécnica Salesiana. <http://dspace.ups.edu.ec/handle/123456789/20696>.
- Barboza, A., & otros, &. (2020). Opinión de la población migrante sobre los cuidados del equipo de salud en dos policlínicas municipales de Montevideo. *Scielo*, 44-53.
- Botella, J., & Zamora, A. (2017). El meta-análisis: una metodología para la investigación en educación. *Educación XXI*, 20(2), 17-38, doi: 10.5944/educXX1.18241.
- Burgo, O. B., León, J. L., Cáceres, M. L., Pérez, C. J., & Espinoza, E. E. (2019). Some thoughts on research and educational intervention. *Revista Cubana de Medicina Militar*, 48(Supl. 1), e383, Recuperado en 29 de julio de 2021, de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0138-65572019000500003&lng=es&tlng=en.
- Cabrera Cárdenas, A. M., & Jiménez Cabrera, K. E. (2017). *PREVALENCIA DE FONEMAS ALTERADOS EN NIÑOS Y NIÑAS DEL DISTRITO*. Obtenido de Repositorio Institucional Universidad de Cuenca:

- <https://dspace.ucuenca.edu.ec/bitstream/123456789/28760/1/PROYECTO%20DE%20INVESTIGACION%20N.pdf>
- Cardoso, C., Talame, L., & Amor, M. (2021). Aprendizaje automático aplicado a la pandemia del virus Covid-19 en Argentina ESPERADOS. *XXIII Workshop de Investigadores en Ciencias de la Computación (WICC 2021, Chilecito, La Rioja)*, 78-81.
- Carrasco-Machado, J., Ramos, A. M., & Morales-Muñoz, R. (2020). Un modelo probabilístico de inventarios por demanda independiente mediante el uso de GRASP. *Ecuadorian Science Journal*, 4(1), 26-31.
- Caturra, R., & Congosto, C. (2016). *Sistema de creación de tutoriales interactivos para el aprendizaje de lenguajes de programación*. Universidad Complutense de Madrid .<https://eprints.ucm.es/id/eprint/38408>.
- Cedeño Pinargote, K. E., & Chinga Quimis, E. A. (2020). *Métricas y pesos para el establecimiento de similitudes en rasgos para el sistema de razonamiento basado en casos, del proyecto FCI para pacientes con Dislalia*. Obtenido de Repositorio Universidad de Guayaquil: <http://repositorio.ug.edu.ec/bitstream/redug/49648/1/B-CISC-PTG-1843-2020%20Cede%20Pinargote%20Katherine%20Edith%20Chinga%20Quimis%20Edisson%20Alberto.pdf>
- Chan May, O. A., Peña Koo, J. J., Vianne Kinani, J. M., & Zapata Encalada, M. A. (2018). Construcción De Un Modelo De Predicción Para Apoyo Al Diagnóstico De Diabetes (Construction of a Prediction Model To Support the Diabetes Diagnosis). *Pistas Educativas*, 2105-2122.
- Chaparro Pesca, J. A., & Cepeda Araque , C. H. (Agosto de 2018). *Repositorio UPTC*. Recuperado el 2021, de Aplicación de un Modelo de Regresión Logística con Respuesta Politémica Nominal en el Análisis de Preferencias Alimentarias de Aves: http://rdigitales.uptc.edu.co/memorias/index.php/apli_estad/apli_esta2016/paper/view/2176
- Chen, J. (2019). A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Elsevier - Environment International*, 1-14.
- Cobo Viera, R. E., & Jaramillo Armijos, B. V. (Febrero de 2020). *Construcción de una base de casos y base transaccional para el sistema del Proyecto FCI Herramienta para el tratamiento de trastornos fonéticos (Dislalia)*. Guayaquil: Repositorio de la Universidad de Guayaquil. <http://repositorio.ug.edu.ec/handle/redug/48980>. Recuperado el Agosto de 2021
- Davidson, L., & González Ibáñez, Á. (2017). La recuperación centrada en la persona y sus implicaciones en salud mental. *Revista de la Asociación Española de Neuropsiquiatría*, 189-205.
- Del Valle Benavides, A. R. (2017). Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones. *Departamento de Estadística e Investigación Operativa*.
- ERP. (2021). *Análisis de requisitos técnicos: procesos y procedimientos*. Obtenido de EVALUANDO ERP.COM: <https://www.evaluandoerp.com/software-erp/conceptos-erp/analisis-requisitos-tecnicos/>
- Escobar Gallegas, J. L., & Bilbao Ramirez, P. H. (2020). *INVESTIGACION Y EDUCACION SUPERIOR*. LULU COM.
- EspacioLogopedico. (2021). *ELCE: Exploracion del lenguaje comprensivo y expresivo*. Obtenido de EspacioLogopedico.com:

- <https://www.espaciologopedico.com/tienda/prod/576/elce-exploracion-del-lenguaje-comprensivo-y-expresivo.html>
- Espinoza, C. (2018). *Diseño y desarrollo de una herramienta lúdica informática para estimular e intervenir los niveles de lenguaje en niños desde los 4 a 5 años*. Cuenca: <https://dspace.ups.edu.ec/>.
- Flores Rodríguez, C. (Marzo de 2021). *Repositorio Institucional Universidad Veracruzana*. Veracruz: Universidad Veracruzana. <https://www.google.com/search?client=firefox-b-d&q=Generaci+%CC%81on+de+%CC%81arboles+ddecisi+%CC%81on+usando+un+algoritmo+inspirado+en+laF+%CC%81%C4%B1sica#>. Recuperado el Julio de 2021, de Generación de árboles de decisión usando un algoritmo inspirado en la Física: <https://www.google.com/search?client=firefox-b-d&q=Generaci+%CC%81on+de+%CC%81arboles+ddecisi+%CC%81on+usando+un+algoritmo+inspirado+en+laF+%CC%81%C4%B1sica#>
- Frascaroli, M. (6 de Junio de 2020). *Facilidad de gestión y rápida implementación: claves en la elección de soluciones de servicio al cliente*. Obtenido de Casos y entrevistas, Chatbots: <https://planetachatbot.com/facilidad-gestion-y-rapida-implementacion-claves-en-eleccion-soluciones-cliente/>
- Fredes Albarracín, E. (2019). *Dislalia: Tipos y tratamientos*. Obtenido de OirPensarHablar: <https://oirpensarhablar.com/dislalia-tipos-tratamiento/>
- Fuentes, M. D., & Medina, W. D. (Abril de 2021). *Diseño de un modelo predictivo-asistencial de pacientes infectados por Covid-19, mediante un modelo supervisado de machine learning basado en criterio de derivacion hospitalaria o ambulatoria*. Obtenido de Repositorio UG: <http://repositorio.ug.edu.ec/bitstream/redug/52650/1/B-CISC-PTG-1901-2021%20Fuentes%20Marmolejo%20Melina%20Daniela%20-%20Medina%20Parra%20Wilmer%20David.pdf>
- Galeano, M. E. (2020). Parte 3. Modalidades de investigación social cualitativa. En M. Galeano, *Diseño de proyectos en la investigación cualitativa* (págs. 13-89). Recuperado de https://books.google.com.ec/books?id=Xkb78OSRMI8C&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false.
- Galvez Quilca, R. (2019). *SISTEMA WEB PARA EL PROCESO DE CONTROL DE PROYECTOS EN LA EMPRESA HUNDRED S.A.C.* César Vallejo: Repositorio Digital Institucional Universidad César Vallejo. <https://hdl.handle.net/20.500.12692/63935>. Obtenido de Repositorio Digital Institucional Universidad César Vallejo.
- Galvez Quilca, R. (2019). *SISTEMA WEB PARA EL PROCESO DE CONTROL DE PROYECTOS EN LA EMPRESA HUNDRED S.A.C.* Obtenido de Repositorio Digital Institucional Universidad César Vallejo: <https://hdl.handle.net/20.500.12692/63935>
- García Girón, E. G. (2020). *Detección de patrones de deserción estudiantil mediante aplicación de Árboles de Decisión C4. 5 en el IESTP "Señor de Chocán" de Querecotillo*. Universidad Cesar Vallejo. <https://hdl.handle.net/20.500.12692/55029>.
- García Torres, M. E., & Espino Quiñones, L. (2018). *Aplicación de minería de datos basado en árboles de decisión para predecir el riesgo de morosidad de los clientes en la empresa de seguros Oncosalud SAC 2018*. Lima: Universidad Autónoma del Perú. <http://repositorio.autonoma.edu.pe/handle/AUTONOMA/700>.

- Geer Mountain Software Corp. (2016). *Stat::Fit*. Obtenido de <https://www.geerms.com/files/114225421.pdf>
- Gilbets, L. (11 de Noviembre de 2020). *Qué es la metodología Kanban y cómo utilizarla*. Obtenido de IEBS: <https://www.iebschool.com/blog/metodologia-kanban-agile-scrum/>
- González González, L. (2005). *Análisis clínico de la comunicación Verbal en niños con dificultades para aprender que cursan el segundo grado en la enseñanza especial*. Obtenido de Repositorio Institucional de la Universidad Central "Marta Abreu" de Las Villas (UCLV): <https://dspace.uclv.edu.cu/bitstream/handle/123456789/4252/Liliana%20Teresita%20Glez%20viejas.pdf?sequence=1&isAllowed=y#:~:text=La%20investigaci%C3%B3n%20del%20lenguaje%20impresivo,1%C3%B3gico%20gramaticales%20y%20fragmentos%20literarios>.
- Guevara Alban, G. P., Verdesoto Arguello, A. E., & Castro Molina, N. E. (2020). Metodologías de investigación educativa (descriptivas, experimentales, participativas, y de investigación-acción). *Revista científica Mundo de la Investigación y el conocimiento*, 163-173.
- Guillén Escamilla, J. E. (Nov de 2020). *Algunas consideraciones sobre el panorama actual de la lingüística clínica*. Obtenido de *Literatura y lingüística* (42) 303-329: <http://dx.doi.org/10.29344/0717621x.42.2598>
- Guzmán Rizo, A. (Mayo de 2020). *Repositorio Institucional de la Universidad de Costa Rica*. Recuperado el Julio de 2021, de Comparación de dos estrategias metodológicas para la estratificación socioeconómica del marco muestral de viviendas de Costa Rica: <http://www.kerwa.ucr.ac.cr/handle/10669/81098?locale-attribute=en>
- Hernández , C. E., & Carpio, N. (2019). Introducción a los tipos de muestreo. *Revista científica del Instituto Nacional de Salud*, 76-79.
- Hernández Escobar, A. A., Ramos Rodríguez, M. P., Placencia López, B. M., Indacochea Ganchozo, B., Quimis Gómez, A. J., & Moreno Ponce, L. A. (2018). *METODOLOGIA DE LA INVESTIGACION CIENTIFICA*. Editorial Área de innovación y desarrollo.
- Hernández Escobar, A. A., Ramos Rodríguez, M. P., Placencia López, B. M., Indacochea Ganchozo, B., Quimis Gómez, A. J., & Moreno Ponce, L. A. (2018). Metodología de la investigación científica. En A. Hernández Escobar, *Capítulo 2: La Ingeniería ambiental como ciencia* (págs. 15-174). doi : <http://dx.doi.org/10.17993/CcyLI.2018.15>.
- INEC. (28 de Noviembre de 2010). *Resultados Provinciales*. Obtenido de INEC: <https://www.ecuadorencifras.gob.ec/wp-content/descargas/Manual/Resultados-provinciales/guayas.pdf>
- Leon Atiquipa, H. E. (Abril de 2018). *Repositorio Institucional de la PUCP*. Recuperado el 2021, de Desarrollo de un modelo algorítmico basado en árboles de decisión para la predicción de la permanencia de un paciente en un proceso psicoterapéutico: <http://repositorio.pucp.edu.pe/index/handle/123456789/149024>
- Lizares Castillo, M. (2017). Comparación de modelos de clasificación: regresión logística y arboles de clasificación para evaluar el rendimiento académico. *UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS*, 1-68.
- Logopedia. (21 de diciembre de 2019). *¿Qué es un Logopeda?* Obtenido de Logopedia y más: <https://www.logopediaymas.es/blog/que-es-un-logopeda/>
- López Herrero, R. (2020). *Propuesta de intervención en un niño con dislalias y TDHA*. Obtenido de Repositorio Documental Universidad de Valladolid:

- <https://uvadoc.uva.es/bitstream/handle/10324/42960/TFG-G4372.pdf?sequence=1&isAllowed=y>
- Lynn Speiser, J., E. Miller, M., Tooze, J., & Lp, E. (November de 2019). A comparison of random forest variable selection methods for classification prediction modeling. *Science Direct*, 93-101. Recuperado el 2021, de A comparison of random forest variable selection methods for classification prediction modeling: <https://www.sciencedirect.com/science/article/abs/pii/S0957417419303574>
- Maillo, J., Luengo, J., Garcia, S., Herrera, F., & Triguero, I. (2018). *Un enfoque aproximado para acelerar el algoritmo de clasificacion Fuzzy kNN para Big Data*. Granada: Universidad de Granada, España. https://sci2s.ugr.es/sites/default/files/ficherosPublicaciones/2514_CAEPIA2018_paper_184.pdf.
- Maisueche Cuadrado, A. (2019). *UTILIZACIÓN DEL MACHINE LEARNING EN LA INDUSTRIA 4.0*. Valladolid: UNIVERSIDAD DE VALLADOLID. <https://core.ac.uk/download/pdf/228074134.pdf>.
- Manrique Rojas, E. (2020). Machine Learning: análisis de lenguajes y herramientas para desarrollo. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 586-599.
- Márquez, F., Ginberg, M., Rivas, M. E., Jarmoluk, V., Bonamico, L., Ron, M., . . . Olmos, L. (2020). Aplicación del aprendizaje automático para predecir la recuperación de pacientes con desordenes prolongados de conciencia. *Congreso Argentino de Informática y Salud*, 228-247.
- Martínez Manzo, R. (2020). *SISTEMA DE MENÚ INTERACTIVO APOYADO POR UN ALGORITMO BASADO EN REGLAS DE ASOCIACIÓN PARA LA ALIMENTACIÓN DE PACIENTES INTERNOS EN UN HOSPITAL. CASO DE ESTUDIO: HOSPITAL DEL PACÍFICO*. ACAPULCO DE JUAREZ: TECNOLÓGICO NACIONAL DE MÉXICO. <https://rinacional.tecnm.mx/jspui/handle/TecNM/1465>.
- Martínez R., J. L. (2019). El proceso de elaboración y validación de un instrumento de medición documental. *Revista Especializada Acción y Reflexión Educativa*, 2.
- Menacho Chiok, C. H. (2017). Predicción del rendimiento académico aplicando técnicas de minería de datos. *Anales Científicos*, 26-33.
- Merino, J. (2017). La potencialidad de la Regresión Logística Multinivel. *Empiria. Revista De metodología De Ciencias Sociales*, (36), 177-211. <https://doi.org/10.5944/empiria.36.2017.17865>.
- Microsoft. (8 de Julio de 2018). *Algoritmo de árboles de decisión de Microsoft*. Obtenido de Microsoft: <https://docs.microsoft.com/es-es/analysis-services/data-mining/microsoft-decision-trees-algorithm?view=asallproducts-allversions>
- Microsoft. (2021). *Actualizaciones Visual Studio Code*. Obtenido de Visual Studio Code : https://code.visualstudio.com/updates/v1_60
- Monja Sandoval, E. A. (2020). *Comparación de naive bayes y series de tiempo en la predicción de morosidad de cuotas sociales del Colegio de Ingenieros del Perú Consejo Departamental Lambayeque*. Pimentel: Universidad Señor de Sipán. <https://hdl.handle.net/20.500.12802/6755>.
- Montoya Zuluaga, C. A. (Julio de 2020). *Competencias ciudadanas para jóvenes en inmersión política*. Pereira: Universidad de Pereira. <https://repositorio.ucp.edu.co/bitstream/10785/6343/1/DDEPDH93.pdf>. Recuperado el Agosto de 2021
- Olmedo Vega, V., Aguilar Idáñez, J., & Arenillas Lara, J. F. (2019). ANÁLISIS DE FACTORES ASOCIADOS A LA RECUPERACIÓN INTEGRAL DE

- PACIENTES DE ICTUS AL ALTA DE HOSPITAL DE AGUDOS. *Revista Española de Salud Pública*, 1-13.
- Origel, C. G., Rendón, E., Abundez, I., & Alejo, R. (2020). *Repositorio Digital Instituto Politécnico Nacional*. Recuperado el 2021, de Redes neuronales artificiales y árboles de decisión para la clasificación con datos categóricos: https://rcs.cic.ipn.mx/2020_149_8/Redes%20neuronales%20artificiales%20y%20arboles%20de%20decision%20para%20la%20clasificacion%20con%20datos%20categoricos.pdf
- Otzen, T., & Manterola, C. (2017). Técnicas de Muestreo sobre una Población a Estudio. *Scielo*, 227-232.
- Pascual Garcia, P. (2012). Parte III. Conceptos y clasificación de las dislalias. En P. Pascual, *La dislalia: naturaleza, diagnóstico y rehabilitación* (pág. 194. Recuperado de <https://elibro.net/es/ereader/uGuayaquil/153571>). España: CEPE.
- Pérez Alonso, D. (Febrero de 2017). *Utilización de Técnicas de Minería de Dato para la Predicción del Comportamiento de Biosólidos Aplicados en Agricultura*. Burgos: UNIVERSIDAD DE BURGOS. doi: <http://hdl.handle.net/10259/4509>. Recuperado el 2021, de Utilización de Técnicas de Minería de Dato para la Predicción del Comportamiento de Biosólidos Aplicados en Agricultura.
- Pérez Díaz, K. N. (2021). PREDICCIÓN DE USO DE UNIDADES DE CUIDADOS INTENSIVOS PARA LA PANDEMIA COVID-19. *UNIVERSIDAD DE CONCEPCIÓN*, 1-87.
- Pérez, M. (11 de febrero de 2021). *Definición de Trastorno*. Obtenido de ConceptoDefinición: <https://conceptodefinicion.de/trastorno/>
- Python. (2020). *Python*. Obtenido de Python Software Foundation: <https://www.python.org/about/>
- Quan, Z., & Valdez, E. A. (July de 2018). *De Gruyter Publishes*. Recuperado el 2021, de Predictive analytics of insurance claims using multivariate decision trees: <https://www.degruyter.com/document/doi/10.1515/demo-2018-0022/html>
- RAE. (2020). Obtenido de Diccionario de la lengua española: <https://dle.rae.es/>
- Raffino, M. (11 de noviembre de 2020). *Fonema*. Obtenido de Concepto.de: <https://concepto.de/fonema/>
- Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística. (2020). : *XXVI Congreso Argentino de Ciencias de la Computación*, 491-500.
- Regalado Macias, R. M., & Calle De la A, J. J. (2020). *Construcción de algoritmo de similitud para el sistema del Proyecto FCI, Herramienta para el tratamiento de trastornos fonéticos (Dislalia)*. Obtenido de Repositorio UG: <http://repositorio.ug.edu.ec/bitstream/redug/48877/1/B-CISC-PTG-1728%20Regalado%20Mac%20c3%20adas%20Rosa%20Mar%20c3%20ada%20Calle%20de%20la%20A%20%20Juan%20Jos%20c3%20a9.pdf>
- Regalado Macias, R. M., & Calle De la A, J. J. (2020). *Construcción de algoritmo de similitud para el sistema del Proyecto FCI, Herramienta para el tratamiento de trastornos fonéticos (Dislalia)*. Guayaquil: Universidad de Guayaquil. <http://repositorio.ug.edu.ec/bitstream/redug/48877/1/B-CISC-PTG-1728%20Regalado%20Mac%20c3%20adas%20Rosa%20Mar%20c3%20ada%20Calle%20de%20la%20A%20%20Juan%20Jos%20c3%20a9.pdf>. Obtenido de Repositorio UG.
- Reyes González, Y., Martínez Sánchez, N., Díaz Sardiñas, A., & Milanés Luque, M. (2017). De la extracción al modelado del conocimiento en un Sistema Basado en el Conocimiento. Un enfoque desde el agrupamiento conceptual lógico

- combinatorio (From the extraction to knowledge modeling in a Knowledge Based System). . *GEOCONTEC: Revista Internacional De Gestión Del Conocimiento Y La Tecnología*, 5 (2), 41-57 Recuperado a partir de: <https://www.upo.es/revistas/index.php/gecontec/article/view/2549>.
- Ríos Sucasaca, K. D. (2017). REGRESIÓN LOGÍSTICA DE LOS FACTORES QUE INFLUYEN EN LA GASTRITIS, HOSPITAL REGIONAL MANUEL NUÑEZ BUITRÓN-PUNO. *UNIVERSIDAD NACIONAL DEL ALTIPLANO*, 1-139.
- Rivas, R. (11 de julio de 2021). *APOYO ACADÉMICO POR ANTOLOGÍAS*. Obtenido de UNAM: <https://www.iztacala.unam.mx/rrivas/diagnostico.html>
- Robles Pastor, B. F. (Febrero de 2019). *Pueblo Continente* . Recuperado el Julio de 2021, de Población o muestra: <http://200.62.226.189/PuebloContinente/article/view/1269/1099>
- Rodríguez, M., & Mendivelso, F. (2018). DISEÑO DE INVESTIGACIÓN DE CORTE TRANSVERSAL. *unisanitas.edu.co*, 1-7.
- Roman, V. (25 de Abril de 2019). *Algoritmos Naive Bayes: Fundamentos e Implementación*. Obtenido de Ciencia & Datos: <https://medium.com/datos-y-ciencia/algoritmos-naive-bayes-fudamentos-e-implementaci%C3%B3n-4bcb24b307f>
- Romero Tabeayo, I., & Rosa N., F. C. (2017). Diseño de un instrumento de observación para la investigación en enseñanza . *REIPE (Revista de Estudios e Investigación en Psicología y Educación)*, 170-173 .
- Rouhiainen, L. (2018). *Inteligencia artificial 101 cosas que debes saber hoy sobre nuestro futuro*. Madrid: Alienta Editorial .
- Sagaró del Campo, N. M., & Zamora Matamoros, L. (2019). Análisis estadístico implicative versus Regresión logística binaria para el estudio de la causalidad en salud. *Scielo*, 1416- 1440.
- Salazar P, C., & Del Castillo G, S. (2018). *Fundamentos básicos de estadística*.
- Salazar, E. J., & Alzate, W. A. (2018). Aplicación de la simulación Monte Carlo en la proyección del estado de resultados. Un estudio de caso. *Espacios 39(51)*, 11 . Recuperado de <http://es.revistaespacios.com/a18v39n51/a18v39n51p11.pdf>.
- Salazar, E. J., & Alzate, W. A. (2018). Aplicación de la simulación Monte Carlo en la proyección del estado de resultados. Un estudio de caso. *Espacios 39(51)*, 11 . Recuperado de <http://es.revistaespacios.com/a18v39n51/a18v39n51p11.pdf>.
- Salesforce Latinoamérica. (22 de junio de 2017). *Inteligencia Artificial: ¿Qué es? - Blog de Salesforce*. Obtenido de salesforce blog: <https://www.salesforce.com/mx/blog/2017/6/Que-es-la-inteligencia-artificial.html>
- Sánchez Pacheco, C. L. (2019). Gamificación: Un nuevo enfoque para la educación ecuatoriana Gamification: A new approach to Ecuadorian education. *Revista Internacional Docente 2.0 Tecnología-Educativa*, 1-10.
- Sánchez Terán, M. d. (2018). *GUÍA DIDÁCTICA DE DETECCIÓN E INTERVENCIÓN DE LA DISLALIA FUNCIONAL EN NIÑOS DE 4 A 6 AÑOS*. Obtenido de Repositorio de la Pontifica Universidad Catolica del Ecuador: <http://repositorio.puce.edu.ec/bitstream/handle/22000/15119/Tesis-Mari%CC%81a%20del%20Carmen%20Sa%CC%81nchez%20Tera%CC%81n%20.pdf?sequence=1&isAllowed=y>
- Sandoval, L. J. (2018). Algoritmos de Aprendizaje Automático para Análisis y Predicción de Datos. *Revista Tecnológica ITCA-FEPADE*, 36-40.
- Schwarz Díaz, M. (2018). *Identificación y caracterización del problema de investigación para la elaboración de la tesis universitaria*. Lima: Repositorio de

- la Universidad de Lima.
<https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwj6xqyRkq3yAhWhSjABHXQqA6UQFnoECAwQAaw&url=https%3A%2F%2Frepositorio.ulima.edu.pe%2Fhandle%2F20.500.12724%2F7099&usg=AOvVaw1QhzkA>. Recuperado el 2021
- Shaier, S. (19 de Febrero de 2019). *ML Algorithms: One SD (σ)- Bayesian Algorithms*. Obtenido de Towards data science: <https://towardsdatascience.com/>
- Shaier, S. (19 de Febrero de 2019). *ML Algorithms: One SD (σ)- Bayesian Algorithms*. Obtenido de Towards data science: <https://towardsdatascience.com/>
- Significados. (11 de Enero de 2017). *Significado de Excel*. Obtenido de Significados: <https://www.significados.com/excel/>
- Solíz Plata, D. J. (2019). *Cómo Hacer Un Perfil Proyecto De Investigación Científica*. Palibrio Recuperado de https://books.google.com.ec/books?id=Q-GCDwAAQBAJ&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false.
- Soni, D. (8 de Junio de 2018). *Introduction to Bayesian Networks*. Obtenido de Toward data science: <https://towardsdatascience.com/introduction-to-bayesian-networks-81031eed94e>
- Sznajdleder, P. (2017). *Algoritmos a fondo: con implementaciones en c y java*. Argentina: Alfaomega.
- Vega Hernández, C. A. (2020). ESTRATIFICACIÓN DE RIESGO DE INFECCIONES BACTERIANAS INVASORAS BASADO EN ALGORITMOS DE MACHINE LEARNING PARA PACIENTES PEDIÁTRICOS INGRESADOS POR NEUTROPENIA FEBRIL. *UNIVERSIDAD DE CHILE FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS DEPARTAMENTO DE INGENIERÍA INDUSTRIAL*, 1-99.
- Vicuña Pinto, A. E., Cortez Chichande , B. S., & Basurto Tovar , Y. J. (2018). Determinación de modelos predictivos para los indicadores de competitividad empresarial aplicando regresión lineal. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 94-107.
- Villabona, J. (4 de Abril de 2018). Base de conocimientos. *¿Cuáles son los requerimientos mínimos que se necesitan en un computador para trabajar exitosamente en el aula virtual?*
- Walker, J. S. (2018). *Python: La Guía Definitiva para Principiantes para Dominar Python*. Babelcube Inc..
- Yang, M.-S., & Sinaga, K. P. (2019). A Feature-Reduction Multi-View k-Means Clustering Algorithm. *IEEE Access*, vol.7, 114472-114486. doi: 10.1109/ACCESS.2019.2934179.
- Yu, S.-S., Chu, S.-W., Wang, C.-M., Chan, Y.-K., & Chang, T.-C. (2018). Two improved k-means algorithms. *ScienceDirect*, 68(), 747-755. doi:10.1016/j.asoc.2017.08.032.