



UNIVERSIDAD AUTÓNOMA  
DE AGUASCALIENTES

CENTRO DE CIENCIAS BÁSICAS  
CIENCIAS DE LA COMPUTACIÓN

**TESIS**

METODOLOGÍA PARA EL ANÁLISIS DE SEÑALES EEG POR MAPEO  
DE REGIONES CEREBRALES PARA LA SALUD MENTAL

**PRESENTA**

JULIA ELIZABETH CALDERÓN REYES

PARA OBTENER EL GRADO DE  
MAESTRA EN CIENCIAS CON OPCIONES A LA COMPUTACIÓN

**TUTORES**

DR. FRANCISCO JAVIER ÁLVAREZ RODRÍGUEZ

DR. HÉCTOR CARDONA REYES

**INTEGRANTE DEL COMITÉ TUTORAL**

MTRA. MARÍA LORENA BARBA GONZÁLEZ

Aguascalientes, Ags, mayo de 2023

**CARTA DE VOTO APROBATORIO  
INDIVIDUAL**

**M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ**  
DECANO (A) DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **TUTOR** designado de la estudiante **JULIA ELIZABETH CALDERÓN REYES** con ID **178522** quien realizó la tesis titulada: **METODOLOGÍA PARA EL ANÁLISIS DE SEÑALES EEG POR MAPEO DE REGIONES CEREBRALES PARA LA SALUD MENTAL**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que ella pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

**ATENTAMENTE**  
**"Se Lumen Proferre"**  
**Aguascalientes, Ags., a día 25 de mayo de 2023.**



**Dr. Francisco Javier Álvarez Rodríguez**  
Co-Tutor de tesis

c.c.p.- Interesado  
c.c.p.- Secretaría Técnica del Programa de Posgrado

**CARTA DE VOTO APROBATORIO  
INDIVIDUAL**

**M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ**  
DECANO (A) DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **CO-TUTOR** designado de la estudiante **JULIA ELIZABETH CALDERÓN REYES** con ID **178522** quien realizó la tesis titulada: **METODOLOGÍA PARA EL ANÁLISIS DE SEÑALES EEG POR MAPEO DE REGIONES CEREBRALES PARA LA SALUD MENTAL**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que ella pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

**ATENTAMENTE**  
**"Se Lumen Proferre"**  
**Aguascalientes, Ags., a día 25 de mayo de 2023.**

*Hector CR.*  
**Dr. Héctor Cardona Reyes**  
**Co-Tutor de tesis**

c.c.p.- Interesado  
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.  
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.  
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-07  
Actualización: 01  
Emisión: 17/05/19

**CARTA DE VOTO APROBATORIO  
INDIVIDUAL**

**M. EN C. JORGE MARTÍN ALFÉREZ CHÁVEZ**  
DECANO (A) DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como **ASESORA** designada de la estudiante **JULIA ELIZABETH CALDERÓN REYES** con ID **178522** quien realizó la tesis titulada: **METODOLOGÍA PARA EL ANÁLISIS DE SEÑALES EEG POR MAPEO DE REGIONES CEREBRALES PARA LA SALUD MENTAL**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que ella pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

**ATENTAMENTE**  
**"Se Lumen Proferre"**  
**Aguascalientes, Ags., a día 25 de mayo de 2023.**



**Mtra. María Lorena Barba González**  
**Asesora de tesis**

c.c.p.- Interesado  
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.  
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.  
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-07  
Actualización: 01  
Emisión: 17/05/19



**DICTAMEN DE LIBERACION ACADEMICA PARA INICIAR LOS TRAMITES DEL EXAMEN DE GRADO**



Fecha de dictaminación dd/mm/aaaa: 25/05/2023

**NOMBRE:** Julia Elizabeth Calderón Reyes **ID** 178522

**PROGRAMA:** Maestría en Ciencias con Opciones a la Computación, Matemáticas Aplicadas **LGAC (del posgrado):** Computación-Ingeniería de Software

**TIPO DE TRABAJO:** ( X ) Tesis ( ) Trabajo Práctico

**TITULO:** Metodología para el Análisis de Señales EEG por Mapeo de Regiones Cerebrales para la Salud Mental

**IMPACTO SOCIAL (señalar el impacto logrado):** Desarrollo metodológico con potencial de coadyuvar al personal del DIF Estatal de Aguascalientes, y especialistas de la salud mental; promoviendo la concientización de los trastornos de ésta, y sustentando la continuidad del proyecto al servicio de la sociedad.

**INDICAR SI NO N.A. (NO APLICA) SEGÚN CORRESPONDA:**

INDICAR	SI	NO	N.A. (NO APLICA)	SEGÚN CORRESPONDA:
<i>Elementos para la revisión académica del trabajo de tesis o trabajo práctico:</i>				
SI				El trabajo es congruente con las LGAC del programa de posgrado
SI				La problemática fue abordada desde un enfoque multidisciplinario
SI				Existe coherencia, continuidad y orden lógico del tema central con cada apartado
SI				Los resultados del trabajo dan respuesta a las preguntas de investigación o a la problemática que aborda
SI				Los resultados presentados en el trabajo son de gran relevancia científica, tecnológica o profesional según el área
SI				El trabajo demuestra más de una aportación original al conocimiento de su área
SI				Las aportaciones responden a los problemas prioritarios del país
NO				Generó transferencia del conocimiento o tecnológica
SI				Cumple con la ética para la investigación (reporte de la herramienta antiplagio)
<i>El egresado cumple con lo siguiente:</i>				
SI				Cumple con lo señalado por el Reglamento General de Docencia
SI				Cumple con los requisitos señalados en el plan de estudios (créditos curriculares, optativos, actividades complementarias, estancia, predoctoral, etc)
SI				Cuenta con los votos aprobatorios del comité tutorial, en caso de los posgrados profesionales si tiene solo tutor podrá liberar solo el tutor
N.A.				Cuenta con la carta de satisfacción del Usuario
SI				Coincide con el título y objetivo registrado
SI				Tiene congruencia con cuerpos académicos
SI				Tiene el CVU del Conacyt actualizado
N.A.				Tiene el artículo aceptado o publicado y cumple con los requisitos institucionales (en caso que proceda)
<i>En caso de Tesis por artículos científicos publicados</i>				
N.A.				Aceptación o Publicación de los artículos según el nivel del programa
N.A.				El estudiante es el primer autor
N.A.				El autor de correspondencia es el Tutor del Núcleo Académico Básico
N.A.				En los artículos se ven reflejados los objetivos de la tesis, ya que son producto de este trabajo de investigación.
N.A.				Los artículos integran los capítulos de la tesis y se presentan en el idioma en que fueron publicados
N.A.				La aceptación o publicación de los artículos en revistas indexadas de alto impacto

Con base a estos criterios, se autoriza se continúen con los trámites de titulación y programación del examen de grado:

Si  X   
No \_\_\_\_\_

**Elaboró:**

\* NOMBRE Y FIRMA DEL CONSEJERO SEGÚN LA LGAC DE ADSCRIPCIÓN:

FIRMAS

Dr. Francisco Javier Álvarez Rodríguez

NOMBRE Y FIRMA DEL SECRETARIO TÉCNICO:

Dr. Hermito Sánchez Cruz

\* En caso de conflicto de intereses, firmará un revisor miembro del NAB de la LGAC correspondiente distinto al tutor o miembro del comité tutorial, asignado por el Decano

**Revisó:**

NOMBRE Y FIRMA DEL SECRETARIO DE INVESTIGACIÓN Y POSGRADO:

Dr. Juan Jauregui Rincón

**Autorizó:**

NOMBRE Y FIRMA DEL DECANO:

M. en C. Jorge Martín Alférez Chávez

**Nota: procede el trámite para el Depto. de Apoyo al Posgrado**

En cumplimiento con el Art. 105C del Reglamento General de Docencia que a la letra señala entre las funciones del Consejo Académico: ... Cuidar la eficiencia terminal del programa de posgrado y el Art. 105F las funciones del Secretario Técnico, llevar el seguimiento de los alumnos.

Elaborado por: D. Apoyo al Posg.  
Revisado por: D. Control Escolar/D. Gestión de Calidad.  
Aprobado por: D. Control Escolar/ D. Apoyo al Posg.

Código: DO-SEE-FO-15.  
Actualización: 01  
Emisión: 28/04/20

## Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología: CONACYT, por la beca para estudios de posgrado de calidad (PNPC) otorgada durante el periodo enero 2021-diciembre 2022; siendo lo que me permitió cursar mis estudios a tiempo completo, dando seguimiento a mi preparación académica.

A mi alma máter, la Benemérita Universidad Autónoma de Aguascalientes, por proyectar en mí la luz del conocimiento a través de las cátedras que recibí en sus instalaciones, dando pauta al desarrollo del proyecto de investigación.

Al DIF Estatal de Aguascalientes por brindar el espacio para las pruebas, contribuyendo a la toma de datos de usuarios para la formulación de casos de estudio y antecedentes.

A mis tutores: el Dr. Francisco Álvarez y el Dr. Héctor Cardona, por su invaluable guía durante el transcurso del posgrado, por cada reto y aprendizaje obtenido que me inspiró a adaptarme y continuar preparándome en el área de las Ciencias Computacionales.

A mi comité tutorial, al Dr. Humberto Muñoz, por compartir su experiencia en las áreas de Inteligencia Artificial y fomentar la integración con Ciencia de Datos, y a la MTF. Lorena Barba por contribuir desde su experiencia en Psicología, brindando un enfoque integral al proyecto de investigación, cimentando así sus bases para las pruebas con usuarios.

A mis docentes en tan maravillosa casa de estudios, por mostrarme las diversas áreas de investigación aplicada de las Ciencias Computacionales y áreas relacionadas, por enseñar con el ejemplo e innovar en técnicas de aprendizaje tras los ajustes por la pandemia.

Al Dr. Marco Hernández y el Dr. Ricardo Mendoza por su guía en etapas tempranas de mi carrera profesional e incursión en la academia, cuyo fomento a la investigación aplicada e invaluable apoyo, me condujeron al presente trabajo de tesis y línea de investigación de la Maestría en Ciencias Computacionales.

A mi colega la MCC. Yuriko Fernández, por compartir conmigo experiencias de aprendizaje inolvidables, por mostrarme que amistad, hermandad y sororidad son sinónimos no solo por escrito sino en la vida real.

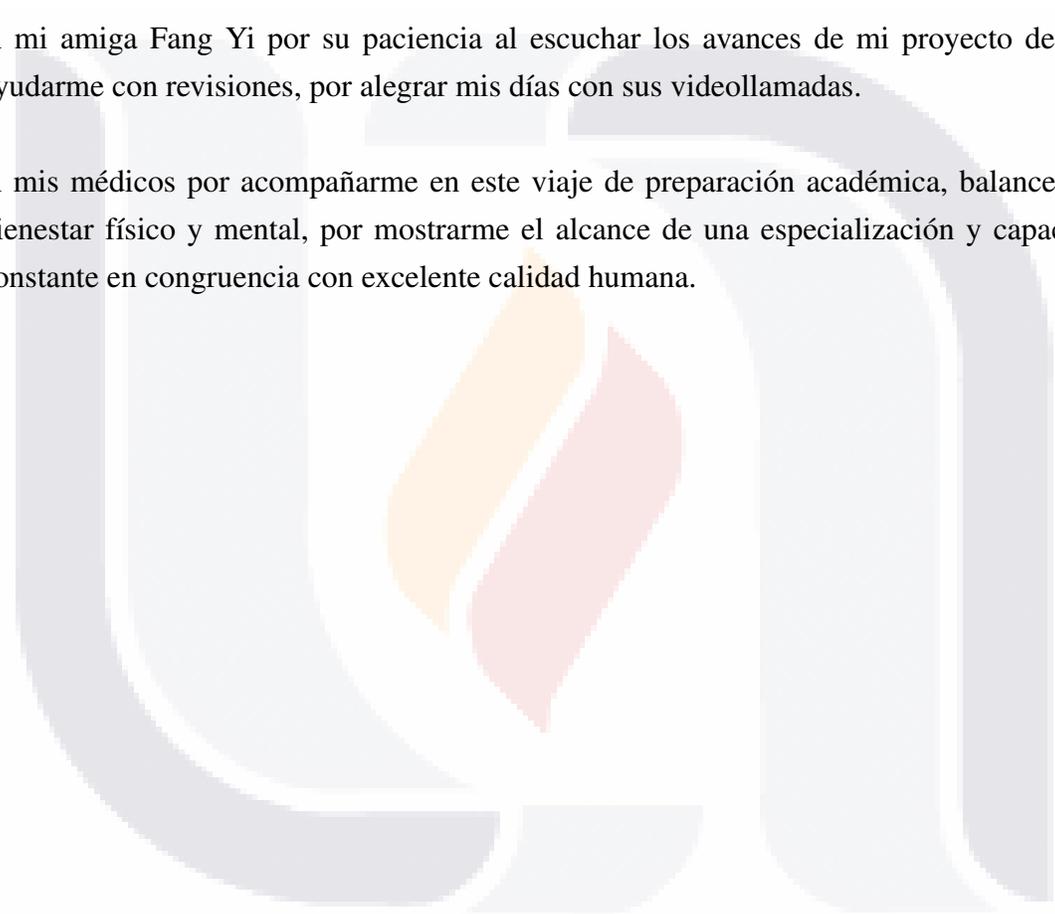
## **Dedicatoria**

A mi madre Rosario, por su apoyo incondicional en mi camino hacia la adquisición y generación del conocimiento; por brindarme su sabiduría y candor en tiempos de cambio.

A mi amiga Crystal por creer en mí y estar presente como la voz de la razón pese a la distancia y el ajetreo de la vida cotidiana, por formar parte de mi vida en éxitos y altibajos.

A mi amiga Fang Yi por su paciencia al escuchar los avances de mi proyecto de tesis y ayudarme con revisiones, por alegrar mis días con sus videollamadas.

A mis médicos por acompañarme en este viaje de preparación académica, balanceando el bienestar físico y mental, por mostrarme el alcance de una especialización y capacitación constante en congruencia con excelente calidad humana.



# Índice General

<b>Índice de Tablas</b>	<b>4</b>
<b>Índice de Figuras</b>	<b>6</b>
<b>Resumen</b>	<b>7</b>
<b>Abstract</b>	<b>8</b>
<b>Introducción</b>	<b>9</b>
<b>1. Generalidades de la Investigación</b>	<b>10</b>
1.1. Planteamiento del Problema . . . . .	11
1.2. Formulación del Problema . . . . .	11
1.3. Propósito y Objetivos . . . . .	12
1.3.1. Propósito . . . . .	12
1.3.2. Objetivo General . . . . .	12
1.3.3. Objetivos Específicos . . . . .	12
<b>2. Marco Teórico</b>	<b>13</b>
2.1. Caracterización del Área . . . . .	13
2.1.1. Ciencias Computacionales: Ingeniería de Software y Ciencia de Datos	14
2.1.2. Ciencias de la Salud: Biomédica y Neurociencia . . . . .	15
2.2. El Estudio del Cuerpo Humano: Aplicaciones de Bioseñales . . . . .	16
2.2.1. Fundamentos de Biomédica . . . . .	17
2.2.2. Clasificación de Bioseñales . . . . .	17
2.2.3. Bioseñales Eléctricas: Electroencefalograma . . . . .	19
2.3. Interacción Humano Computadora e Interfaces . . . . .	20
2.3.1. Fundamentos de Interacción Humano Computadora . . . . .	20

2.3.2.	Interfaces Cerebro-Computadora . . . . .	21
2.3.3.	Representación del Conocimiento . . . . .	22
2.4.	Comportamiento Humano: Salud Mental y Afecciones . . . . .	23
2.4.1.	La Conducta Humana e Influencia de los Modelos Psicológicos . . . . .	24
2.4.2.	Caracterización de Trastornos Mentales del Estado de Ánimo . . . . .	25
2.4.3.	Representación de la Ansiedad y sus Factores Clave . . . . .	26
2.5.	Metodología de Ciencia de Datos: Aplicaciones de IA . . . . .	28
2.5.1.	Metodología en el Marco de Ingeniería de Software . . . . .	29
2.5.2.	Ciencia de Datos para la Resolución de Problemas . . . . .	30
2.5.3.	Técnicas Algorítmicas: Clasificación y Categorización . . . . .	31
2.5.4.	Comparación de Metodologías y Caracterización . . . . .	36
<b>3.</b>	<b>Metodología</b>	<b>39</b>
3.1.	Propuesta Metodológica . . . . .	41
3.1.1.	Segmentación: Tamaño de la Muestra y Enfoque . . . . .	42
3.1.2.	Diseño: Esquemmatización y Prototipado . . . . .	42
3.1.3.	Tratamiento: Análisis y Modelado . . . . .	43
3.1.4.	Lanzamiento: Experimentación y Recolección . . . . .	43
3.1.5.	Evaluación: Simulación y <i>Feedback</i> . . . . .	44
3.2.	Guía de Diseño de <i>MVPs</i> . . . . .	44
3.2.1.	Segmentación: Enfoque del Equipo Multidisciplinario . . . . .	45
3.2.2.	Diseño: Esquema del Sistema a Prototipar . . . . .	45
3.2.3.	Tratamiento: Análisis y Modelado de Datos . . . . .	46
3.2.4.	Lanzamiento: Pruebas de uso de la <i>BCI</i> . . . . .	47
3.2.5.	Evaluación: Simulación de Datos Locales . . . . .	48
3.3.	Diseño Experimental . . . . .	49
3.3.1.	Segmentación de los Datos: Recolección, Exploración y Limpieza . . . . .	50
3.3.2.	Abstracción de los Datos: Análisis, Tratamiento y Estructura . . . . .	50
3.3.3.	Modelado de Datos: Muestreo, Entrenamiento e Implementación . . . . .	51
3.3.4.	Perfilado de Datos: Evaluación, Optimización y Comparación . . . . .	52
<b>4.</b>	<b>Resultados</b>	<b>55</b>
4.1.	Contextualización Instrumental . . . . .	55
4.1.1.	Implementación del Algoritmo K-Means . . . . .	55
4.1.2.	Implementación de los Algoritmos ID3 y J48 . . . . .	57
4.2.	Asignación de Procesos de IA . . . . .	59

**ÍNDICE GENERAL** 3

4.2.1. Segmentación: Variables de Estudio Globales . . . . . 60

4.2.2. Abstracción: Transformación del Espacio Experimental . . . . . 61

4.2.3. Modelado de Datos: Implementación del Modelo de IA . . . . . 62

4.2.4. Perfilado: Evaluación e Iteración del Modelo . . . . . 63

**5. Conclusiones** 65

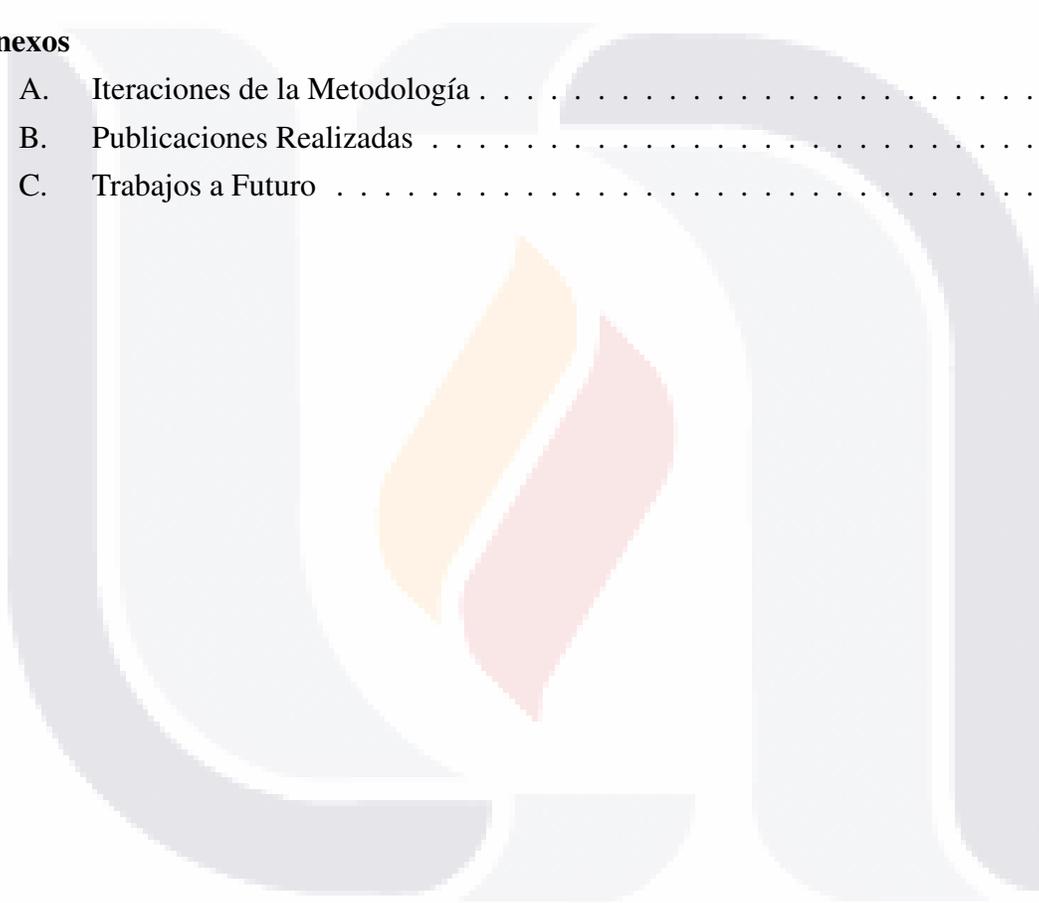
**Referencias** 68

**Anexos** 87

A. Iteraciones de la Metodología . . . . . 88

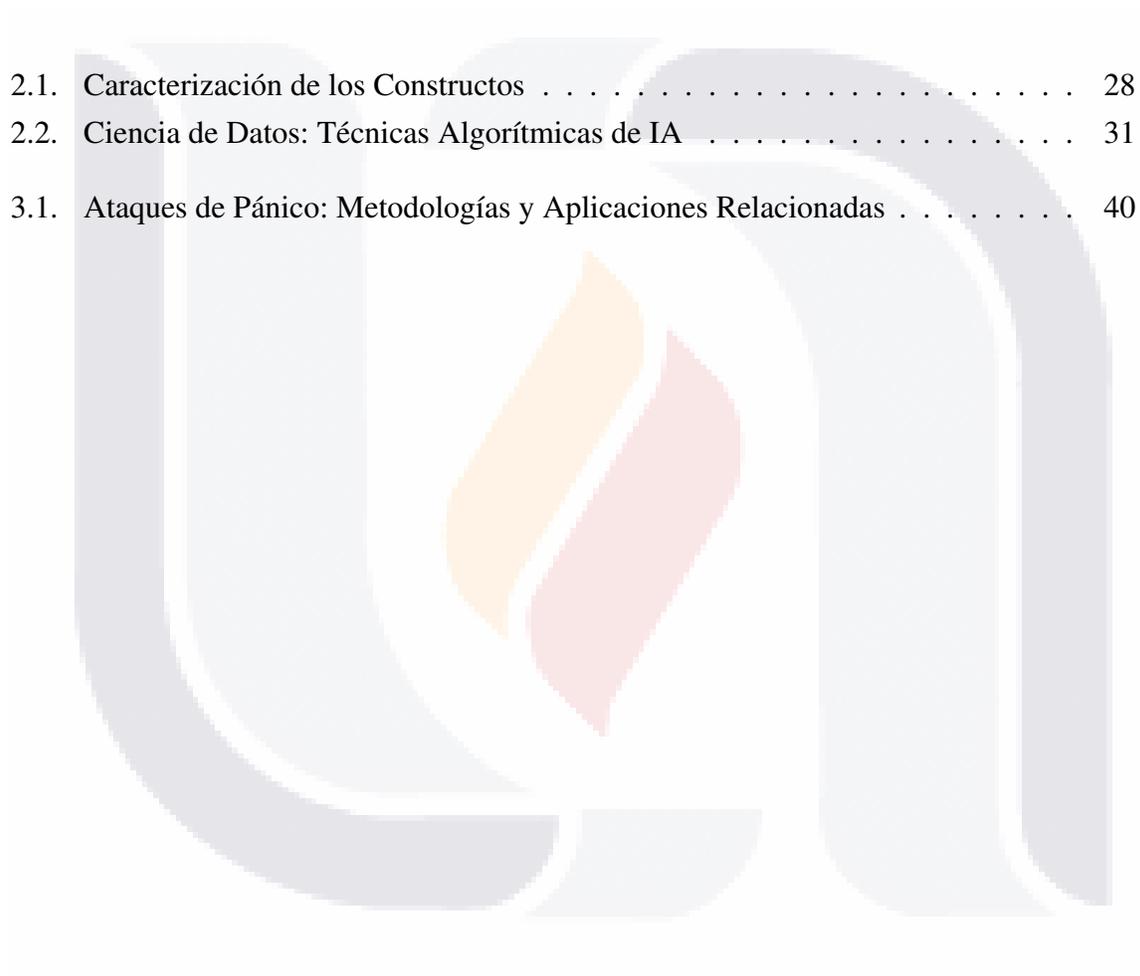
B. Publicaciones Realizadas . . . . . 90

C. Trabajos a Futuro . . . . . 93



## Índice de Tablas

2.1. Caracterización de los Constructos . . . . .	28
2.2. Ciencia de Datos: Técnicas Algorítmicas de IA . . . . .	31
3.1. Ataques de Pánico: Metodologías y Aplicaciones Relacionadas . . . . .	40



# Índice de Figuras

2.1.	Caracterización del Área. . . . .	13
2.2.	Pseudocódigo para el Algoritmo de <i>K-Means</i> . . . . .	33
2.3.	Pseudocódigo para el Algoritmo <i>ID3</i> . . . . .	35
2.4.	Ciclo de la Metodología de <i>Lean UX</i> [56]. . . . .	37
2.5.	Flujo de la Metodología de Ciencia de Datos de <i>IBM</i> [118]. . . . .	38
3.1.	Desglose del Diseño Metodológico [31] Inspirado en <i>Lean UX</i> [56]. . . . .	41
3.2.	Esquema del Equipo Multidisciplinario. . . . .	45
3.3.	Diagrama de Clases del Sistema. . . . .	46
3.4.	Diagrama de Contexto del Sistema. . . . .	47
3.5.	Diagrama de Contenedores del Sistema. . . . .	48
3.6.	Evaluación del Sistema. . . . .	48
3.7.	Contextualización de los Procesos de Ciencia de Datos e IA [30]. . . . .	49
4.1.	Determinación de <i>K</i> con el Método de Codo. . . . .	56
4.2.	Dispersión del Conjunto de Datos y Clusterización. . . . .	57
4.3.	Árbol de Clasificación <i>ID3</i> . . . . .	58
4.4.	Matrices de Confusión de <i>ID3</i> y <i>J48</i> . . . . .	59
4.5.	Análisis Exploratorio del Modelo. . . . .	60
4.6.	Variables de Estudio del Conjunto de Datos. . . . .	61
4.7.	<i>Script</i> de Tratamiento de Datos en <i>R</i> . . . . .	62
4.8.	<i>Script</i> de Reestructuración de Datos en <i>R</i> . . . . .	63
4.9.	Diseño del Proceso Instrumental. . . . .	64
A.1.	Desglose de Iteraciones de la Metodología. . . . .	89
B.1.	Constancia de Presentación por Artículo en el <i>HCI 2022</i> . . . . .	90
B.2.	Constancia de Presentación de Artículo en el <i>CIMPS 2022</i> . . . . .	91
B.3.	Constancia de Primer Lugar en el <i>MICAI 2022</i> . . . . .	92

C.1. Proyección de Trabajos a Futuro. . . . . 93



## Resumen

Mientras que las bioseñales *EEG* proveen un nivel de entendimiento respecto a la actividad cerebral, métodos computacionales de Ingeniería de Software e Inteligencia Artificial pueden contribuir al desarrollo de un amplio espectro de soluciones. Esta investigación aborda la dificultad para la obtención de datos en el monitoreo de la actividad cerebral al momento de un ataque de pánico, en conjunto con la carencia metodologías y modelos enfocados a trastornos mentales, particularmente ansiedad, lo cual puede contribuir al desarrollo de soluciones basadas en el nivel de ansiedad del individuo y condición en general, por tanto, optimizando procesos basados en datos médicos; se propone un enfoque integral de Ciencia de Datos e Ingeniería de Software como contribución al desarrollo de soluciones orientadas a datos y la obtención parcial de resultados listos para su implementación, ello dentro del alcance metodológico de *Lean UX* y sus *MVPs* para la validación y verificación transversal de la transparencia de los datos de historia clínica de personas con ansiedad, haciendo uso de un dispositivo para la lectura de bioseñales *EEG*, y priorizando la identificación de directrices que permitan el análisis de los datos monitoreados y factores de correlación en personas con ansiedad e historia de ataques de pánico, cimentando la optimización de procesos para especialistas de la salud mental y las sesiones que sostienen con los pacientes.

La propuesta introducida en estas páginas yace en una metodología holística cuyos ejes centrales de *Lean UX* y la Metodología de Ciencia de Datos de *IBM* sustentan el diseño experimental de un modelo subsecuente a la propuesta metodológica, proveyendo la base para la instrumentación y recolección de datos de usuarios objetivo mediante un análisis preliminar; a través de técnicas algorítmicas de clasificación y agrupación que fueron analizadas y comparadas con base a su apego a los requerimientos y restricciones del proyecto de investigación; variables de control fueron identificadas para la mejora en la calidad de los datos, acentuando los resultados obtenidos mediante la comparación entre las técnicas algorítmicas de clusterización con *K-Means*, y árboles de decisiones con *ID3* y *J48*.

**Palabras Clave:** Ciencia de Datos, *Lean UX*, Diseño Metodológico, Bioseñales *EEG*, Ansiedad.

## Abstract

While EEG signals provide insight into brain activity, computational methods from software engineering and artificial intelligence can contribute to the development of a wide range of solutions. This research addresses the difficulty of data acquisition regarding the monitoring of brain activity at the time of a panic attack, coupled with the lack of methodologies and models oriented to mental disorders particularly anxiety, which could contribute to the development of solutions based in the level of anxiety of the user and overall condition, therefore optimizing the processes; an approach encompassing data science and software engineering is proposed to contribute with the development of data-driven solutions and partial acquisition of results ready to be deployed within the means of the Lean UX Methodology MVPs to cross validate the data transparency of clinical history of people with anxiety making use of an EEG device, thus prioritizing the identification of guidelines that allow the analysis of monitored data and correlating factors in people with anxiety and history of panic attacks, paving the way for the health specialist to optimize processes in the sessions held with people that have had anxiety or has anxiety.

The proposal introduced in these pages covers an experimental design of a model based on a holistic methodology with the core principles of Lean UX and the IBM Data Science Methodology. To provide the basis for the instrumentation and recollect data from the target users a preliminary analysis was performed, through algorithmic techniques for classification and clustering that were analyzed and compared in regard to their optimality for the project requirements and constraints; control variables were identified to improve the data quality, hence highlighting the results obtained via comparison between K-Means Clustering and ID3/J48 decision trees.

**Keywords: Data Science, Lean UX, Methodology Design, EEG Biosignals, and Anxiety.**

## Introducción

La Ingeniería de Software se ha consolidado como un área clave para el desarrollo de *frameworks* de integración de servicios basados en datos provistos por el usuario [101], por tanto, estableciendo un nexo entre volúmenes de datos y servicios a través de estatutos presentes en: metodologías, modelos, y procesos; aunque la base para el flujo de los datos y la retroalimentación obtenida para mejorar la calidad del servicio son prioritarios para el desarrollo e implementación de productos de *software*, dada la función central de la interacción humana y sus métricas de *User Experience (UX)* y *User Interface(UI)*, no solo se requiere que el flujo de datos sea evaluado respecto a su accesibilidad y disponibilidad, sino considerado como punto de origen para moldear la percepción adquirida en etapas tempranas de adquisición y tratamiento de datos, por tanto, aplicando principios de Ciencia de Datos para sus procesos.

El presente documento abarca el planteamiento sobre el cual se formuló la metodología propuesta, el desarrollo de esta, y el inherente proceso de diseño para el modelo experimental de correlación entre bioseñales *EEG* y mapeo de regiones cerebrales en el estudio de ansiedad y ataques de pánico, mediante la integración de áreas de estudio primarias de las Ciencias Computacionales como lo son la Ingeniería de Software y la Ciencia de Datos, así como las ramas de *Human Computer Interaction (HCI)* y el enfoque provisto por *Lean UX* [56] que prioriza los principios de *Brain Computer Interfaces (BCI)* y la obtención de resultados parciales en cada *Minimum Viable Product (MVP)* para el análisis y tratamiento de los datos obtenidos mediante casos de estudio de ansiedad en presencia de ataques de pánico en momentos dados de la vida de una persona [75]; proveyendo directrices para su monitoreo de forma cotidiana, al abarcar la identificación de áreas de oportunidad y la optimización de soluciones existentes basadas en el uso de bioseñales *EEG* de pacientes con ansiedad, incluyendo el tratamiento, historia clínica, y diagnóstico de trastorno de ansiedad (particularmente ataques de pánico), para determinar factores de correlación entre la actividad cerebral monitoreada y las técnicas estadísticas orientadas a la Ciencia de Datos con el objetivo de monitorear los constructos de ansiedad y su efecto, partiendo del enfoque de la Ingeniería de Software como un *framework* para sustentar la replicación de las etapas de investigación, así como la canalización de los algoritmos de Inteligencia Artificial (IA) y Ciencia de Datos como procesos clave para la validación cruzada de datos analizados con el dispositivo *EEG* y el diseño del instrumento para medir la respuesta del usuario durante la sesión, presentando los resultados por etapa e iteración como *MVPs* a ser interpretados como indicadores en un *dashboard* para futuras etapas de desarrollo.

# Capítulo 1

## Generalidades de la Investigación

La vigente brecha científica y tecnológica en la salud mental debido a la estigmatización de sus trastornos se ve reflejada en las discrepancias entre el reconocimiento del trastorno, el acceso a los servicios médicos correspondientes, tratamiento y perspectiva de profesionales de la salud [97, 92]; aunado al evento de aislamiento por COVID-19, en el marco del cual en la literatura se han identificado respuestas emocionales a la situación de estrés desencadenada, la presencia de síntomas de ansiedad, presencia u descontrol del trastorno de ansiedad ante crisis [100], así como el impacto a la población de adolescentes y adultos jóvenes que desarrolló crisis socio-emocionales, de ansiedad, y trastornos [32], donde con objeto de la investigación se dio prioridad a la afectación de los ataques de pánico, volviéndose no solo un referente epidemiológico, sino un factor de afectación del bienestar socio-emocional [103, 37]; se propone priorizar la actividad cerebral ante ansiedad y ataques de pánico para la generación de un mapa de la actividad *EEG* registrada y asociación de los *triggers* durante el evento, lo que posteriormente puede utilizarse para aplicar la metodología y el modelo generalizado de correlación entre eventos en diseños experimentales y casos de estudio como parte de un proceso para el desarrollo de tecnologías basadas en el usuario que puedan proporcionar conocimiento con base al uso de bioseñales *EEG*.

La presente investigación consiste en el estudio y mapeo de regiones cerebrales por actividad *EEG* de ansiedad y ataques de pánico, el análisis de cambios en marcadores biológicos asociados a la ansiedad como epinefrina, cortisol, y dopamina ante estímulo externo [46, 119]; el comportamiento de la amígdala [137], e identificación de *triggers* como factor de correlación entre la población muestra durante el monitoreo por *EEG* en la aplicación instrumental, la cual surgió al detectarse un área de oportunidad en el uso de técnicas de las Ciencias Computacionales en neuroimágenes para representar los hallazgos asociados a la psicopatología [48] de ansiedad y ataques de pánico; su alcance yace en el monitoreo de

actividad cerebral para trastornos de ansiedad, en específico de ataques de pánico en una población muestra de 15 a 40 años, para definir factores de correlación, *triggers* y mapear el comportamiento observado. Destaca la aplicación de una metodología basada en Ciencia de Datos y Programación Orientada a Componentes para el procesamiento de datos [13, 135], mediante la aplicación de un modelo basado en factores de correlación entre indicadores clave de la actividad observada como alternativa al monitoreo de un ataque de pánico.

## 1.1. Planteamiento del Problema

La actividad cerebral puede ser difícil de monitorear al momento de un ataque de pánico, ya que requiere de un equipo específico acorde a las partes del cerebro por monitorear y el objetivo de su estudio, tomando en cuenta *EEG* y *fMRI*, adicionalmente factores como estímulos internos y externos pueden afectar la respuesta de la amígdala, influenciar el rango de síntomas a presentar y por ende el reflejo fisiológico a nivel químico [49, 145, 114], siendo que las hormonas involucradas en la química cerebral para ataques de pánico y trastornos de pánico pueden diferir entre sí, de forma que no se presente asociación de una de las hormonas más importantes para el sistema de recompensa del cerebro como es la serotonina, o bien que exista la presencia de colecistoquinina u otras sustancias asociadas [26, 115]; si bien se encuentran en la literatura estudios de *EEG* para el monitoreo de trastornos de ansiedad y ataques de pánico, las herramientas de diagnóstico u revisiones a tratamientos son el común denominador [42, 148], más no una metodología que establezca factores de correlación en la actividad cerebral observada como un modelo generalizable para estos. El presente estudio cuantitativo propone el mapeo de la actividad cerebral como marco de comparación, considerando los *triggers* para una población muestra de personas con trastorno de ansiedad y ataques de pánico de entre 15 y 40 años.

## 1.2. Formulación del Problema

**Hipótesis.** La metodología propuesta de Ciencia de Datos provee un marco de referencia para la generación y transferencia del conocimiento mediante el análisis de biobancos de datos y estudio etnográfico de personas con ansiedad, coadyuvando al proceso de canalización, monitoreo y tratamiento de trastornos de ansiedad a la par de adquirir información de ataques de pánico y sus *triggers*.

### 1.3. Propósito y Objetivos

A continuación, se muestran los propósitos, objetivo general, y objetivos específicos:

#### 1.3.1. Propósito

Actualmente, no se cuenta con estudios en la literatura enfocados a la actividad cerebral antes y después de ataques de pánico para reportar la correlación entre el evento y cambios en la actividad monitoreada, así como los *triggers* de la población muestra. El propósito de esta investigación es analizar bioseñales *EEG* de pacientes diagnosticados con trastorno de ansiedad y ataques de pánico, para el mapeo de regiones cerebrales involucradas; tomando como base técnicas estadísticas y metodologías orientadas a Ciencia de Datos para determinar el factor de correlación existente entre la actividad cerebral de ansiedad y ataque de pánico, considerando los *triggers*.

#### 1.3.2. Objetivo General

Diseñar una metodología de bioseñales *EEG* de pacientes que presenten ansiedad para formular el factor de correlación existente entre la actividad cerebral en crisis de ansiedad, trastornos de ansiedad y ataques de pánico con los estímulos y actividad monitoreada, tomando como base técnicas estadísticas orientadas a la ciencia de datos para representar la categorización y clasificación de los constructos asociados.

#### 1.3.3. Objetivos Específicos

1. Sustentar la propuesta metodológica mediante las proyecciones por etapas e iteración, así como los *Minimum Viable Products (MVPs)*.
2. Asociar los *MVPs* al nivel de los componentes de software y el enfoque de la Ciencia de Datos para resolución de problemas.
3. Diseñar el modelo de IA con base a los procesos para la extracción y clasificación de bioseñales *EEG*.
4. Comparar algoritmos de clasificación y agrupación de IA para su posterior implementación en experimentos con usuarios.
5. Aplicar el modelo de IA al conjunto de datos para la caracterización de parámetros de ansiedad.

# Capítulo 2

## Marco Teórico

### 2.1. Caracterización del Área

El presente proyecto se desarrolló en la Benemérita Universidad Autónoma de Aguascalientes, tomando como base la interconexión entre el marco de la Ingeniería de Software, las técnicas de Ciencia de datos como parte de las Ciencias Computacionales, y contexto de aplicación orientado a las Ciencias de la salud mediante la Biomédica y Neurociencia. Ello, ejemplificado en la Figura 2.1 de caracterización del área y contexto de aplicación:



Figura 2.1: Caracterización del Área.

En que, la Ciencia de Datos funge como la integración de métodos algorítmicos de IA para el estudio y análisis de bioseñales *EEG* en conjunto con el enfoque y herramientas provistas por las Ciencias de la Salud: Biomédica y Neurociencia; englobando en Ciencias de la Computación los datos e información obtenida, así como procesos implementados mediante el *framework* de la Ingeniería de Software, dando cierre a la capa de aplicación, enlazando los hallazgos provistos por la Ciencia de Datos, y fundamentando la línea de investigación del posgrado en cuestión para la investigación.

1. **Ciencia de Datos.** Aplicación de la metodología de Ciencia de Datos, análisis estadístico, optimización de algoritmos, modelado y técnicas de procesamiento, clasificación de la información, y visualización de resultados obtenidos.
2. **Biomédica.** Lectura y análisis de bioseñales, en específico de ondas cerebrales, *EEG*, sus características, e instrumentos de medición disponibles.
3. **Neurociencia.** Marcadores neurológicos, hormonas asociadas a la presencia o ausencia de ansiedad, comportamiento del hipocampo, amígdala y áreas activas del cerebro humano ante situaciones de angustia.
4. **Ingeniería de Software.** Nexos entre las áreas previamente mencionadas mediante el diseño de metodologías y optimización de procesos, partiendo de *HCI*, *UX*, y *UI* a la programación basada en componentes.

### 2.1.1. Ciencias Computacionales: Ingeniería de Software y Ciencia de Datos

La Ingeniería de Software es un área primordial de las Ciencias Computacionales, que moldea soluciones en instancias documentables y esquemas que transforman un prototipo en un ambiente funcional, monitoreando las interacciones de diversos tipos de usuarios para tomar decisiones precisas respecto a la mejora continua implicada en el ciclo de desarrollo; a su vez, la implementación de metodologías de *software* permite optimizar, y regular dicho proceso, puesto que independientemente del área de aplicación las metodologías se encuentran diseñadas para satisfacer las necesidades y requerimientos fijados por el usuario en conjunto con el equipo de trabajo. La gestión de sus prácticas es referente documental para la implementación de procesos y diseño de modelos con diversos enfoques como lo son *Lean* y *Agile*, complementando alternativas existentes en la academia y en la industria [88], las

cuales pueden abordar el perfil de *HCI* priorizando la importancia del usuario en la investigación como parte de la observación y análisis efectuados [104]; subsecuente a la viabilidad de adaptar técnicas de la Ingeniería del Conocimiento como las ontologías de para dar solución a problemas de desarrollo de *software* [34], lo que sustenta la implementación de nuevas tecnologías y *Transfer Learning*, destacando el uso particular de datos biométricos como el pilar del desarrollo de *software* y mejores prácticas [93], al igual que la optimización de procesos e implementación de técnicas algorítmicas de *Machine Learning* y *Deep Learning* [10].

Por otra parte, la Ciencia de Datos es un área derivada de las Ciencias Computacionales que ha crecido exponencialmente dado el volumen de datos y campos de aplicación, suscitando el planteamiento e implementación de metodologías con el enfoque de Ciencia de Datos [82] para solventar objetivos específicos a la interconexión de dispositivos y el origen de los datos, especificaciones que presentan un desafío al momento de formular métricas analíticas, y la aplicación de técnicas existentes [91]; matemáticamente hablando, una de las áreas con mayor demanda de metodologías de Ciencia de Datos parte de la estadística [17, 64] dada la aplicación de conocimiento previo, la obtención de hallazgos para la creación de indicadores, así como aplicación de técnicas que cubren aspectos básicos y complejos. Por lo anterior, aunque un enfoque metodológico forma parte de la base de todo proyecto de *software*, su complemento se encuentra en la tecnología propicia para ello, a lo cual para *HCI* los principios de *UX* y *UI* se ven cubiertos para los fines del proyecto de investigación mediante la implementación de *BCI* que pueden resolver problemas orientados a Tecnologías de la Información, Robótica, IA, Ciencias Cognitivas e inclusive Psicología [72]; a su vez, dentro del marco de IA y técnicas aplicadas en Ciencia de Datos, se endorsa el método científico y los requerimientos, fundamentando el enfoque cualitativo [69] y epistemológico [54] para la implementación de soluciones de *software* basado en datos.

### 2.1.2. Ciencias de la Salud: Biomédica y Neurociencia

El nexo entre las Ciencias Computacionales y las Ciencias de la Salud para el área de aplicación en cuestión, se constituyó tomando como base las áreas de Biomédica y Neurociencia; dado el rango de aplicaciones modernas para electroencefalografía y frecuencias de onda provistas por bioseñales *EEG* a través de dispositivos móviles que son más accesibles para investigadores y público en general de forma que puedan adquirir y hacer uso de estos [77]. Tomando el cuidado de la salud como referente del uso de los sistemas computacionales para los procesos de adquisición, almacenamiento y visualización de datos en el desarrollo de productos de *software* [141] mediante la integración de aprendizaje virtual y dispositivos mé-

dicos, lo cual promueve la implementación de *frameworks* orientados a la transferencia del conocimiento [87] y su gestión pragmática [96].

Por consiguiente, el enfoque de Neurociencia permite englobar el análisis de patrones temporales y espaciales mediante el uso del electroencefalograma y lecturas *EEG* obtenidas acorde al método y paradigma [52], reforzando las técnicas algorítmicas de IA y Ciencia de Datos como parte del enfoque cualitativo en áreas como Psicología y Psiquiatría para el estudio de la salud mental, particularmente como caso de estudio de los trastornos de ansiedad y ataques de pánico, mediante el referente biológico que provee en conjunto con la identificación de regiones cerebrales de la actividad monitoreada; como ejemplo de ello, la posibilidad de formular un sistema para la clasificación de emociones y sentimientos, o bien la presencia de ansiedad e identificación de sus estados con base a un estímulo fisiológico aplicable al seguimiento en sesiones de terapia psicológica [98], consolidando el uso de bioseñales *EEG* ante la vertiente de interacción provista por la caracterización del área de aplicación y el análisis de estas desde *HCI* como enfoque amigable hacia los usuarios [28].

## 2.2. El Estudio del Cuerpo Humano: Aplicaciones de Bioseñales

Si bien la percepción del ser humano puede verse bajo la influencia del entorno, es mediante la conjunción entre la observación e introspectiva que se han generado contribuciones en las áreas de estudio respecto de este, dadas las funciones del cuerpo humano que permiten observar, analizar y recolectar información así como censar lo que le rodea en indicadores de temperatura e incluso estados de ánimo para actuar acorde; lo anterior, con relación a la aplicación de ingeniería inversa mediante el estudio del entorno como un modelo de la naturaleza para generar soluciones biomiméticas [83], o biométricas para la identificación de rasgos característicos de una persona que por sí mismos o en conjunto puedan proveer una forma de autenticación con el ínfimo grado de error [108].

Lo anterior, abordado particularmente desde la Biomédica y sus aplicaciones para recolectar, analizar y procesar datos biológicos que permitan determinar el estado en que se encuentra una persona, o bien su predisposición genética, generando bases de datos biológicas [15] que con fundamento ético y un procesamiento inteligente pueden llegar a convertirse en una de las principales fuentes de información para la generación del conocimiento y validación de transferencia tecnológica en el sector salud para el monitoreo de los pacientes y soluciones de procesamiento [143].

### 2.2.1. Fundamentos de Biomédica

La biomédica es la piedra angular bajo la cual aplicaciones médicas e intervenciones científicas y tecnológicas en el sector salud se moldean acorde a datos biológicos proporcionados por bases de conocimiento; cuenta con técnicas e instrumentación especializada para realizar mediciones e interpretar posibles diagnósticos acorde al uso de los datos [3], hablándose en el caso de la IA, de algoritmos bioinspirados que dan solución a problemas complejos replicando el comportamiento de especies del reino animal, o algoritmos biomiméticos que replican el comportamiento de especies de flora y fauna para generar estructuras con materiales a escala micro y nano que puedan ser adaptables al cuerpo humano [68].

Debido al extenso rango de aplicaciones y sensibilidad de la información generada por el tipo de datos origen y el proceso de recolección de estos, resulta inherente la inclusión de la ética como parte central de los protocolos para el desarrollo de herramientas que tengan como origen de datos bioseñales, puesto que al igual que los profesionales de la salud cuentan con un código de ética así como estatutos para cuidar la privacidad e integridad de la información así como de sus pacientes, las intervenciones científicas y tecnológicas se ven reguladas bajo documentos normativos [59] que respalden la transparencia de resultados y salvaguarden los intereses de las partes involucradas; partiendo de la formación de comités de bioética como medios de regulación y validación de procesos, así como intervención práctica en la formación profesional [116].

### 2.2.2. Clasificación de Bioseñales

Las funciones que realiza el cuerpo humano pueden ser monitoreadas en término de las señales eléctricas que envía el cerebro a los órganos o partes del cuerpo, así como sonidos o pulsaciones que emite, englobando fenómenos fisiológicos que pueden ser observados o censados para llevar a cabo un análisis estadístico y proporcionar un diagnóstico congruente con monitoreo, terapia y/o tratamiento.

De igual manera, las bioseñales se caracterizan en función del método para su clasificación [71] tal y como se describe a continuación:

- **Método basado en Existencia.** Evaluación del estado natural o artificial bajo el cual se realiza la medición de las bioseñales.
  - Bioseñales Permanentes. De estado presente constante en el cuerpo humano sin algún estímulo externo.

- Bioseñales Inducidas. De estado observable ante estimulación externa, y que en su ausencia se vuelve imperceptible para medición.
- **Método basado en Dinamismo.** Evaluación cuantitativa ante la presencia y ausencia de estímulo, así como el estado de la bioseñal.
  - Bioseñales Cuasiestáticas. Presencia de actividad a un nivel estable con cambios ínfimos en un periodo.
  - Bioseñales Variantes. Presencia de cambios extensivos en un periodo, considerablemente notables.
- **Método basado en Origen.** Caracterización particular de las señales por criterios de agrupación cuál taxonomía.
  - Bioseñales Eléctricas. Presencia de actividad eléctrica en neuronas y músculos.
  - Bioseñales Magnéticas. Resultado de la inducción de un campo magnético para lecturas ante estímulos externos.
  - Bioseñales Mecánicas. Deformaciones del cuerpo humano, o bien vibraciones de la piel.
  - Bioseñales Ópticas. Medición de la absorción y dispersión de la luz como cambios en su propagación.
  - Bioseñales Acústicas. Evaluación de sonidos corporales perceptibles con y sin estímulo externo.
  - Bioseñales Químicas. Resultados de composición química en estado líquido, sólido y gaseoso como cambios temporales.
  - Bioseñales Térmicas. Evaluación de los mecanismos heterogéneos para la pérdida y absorción de calor del cuerpo humano.

Aunque no existe una clasificación universal de carácter rígido, se conserva la naturaleza del evento para la obtención de bioseñales, y enfoque cuantitativo para medición y análisis estadístico, que con base a herramientas de modelado puede desglosar su segmentación, tendencias y posible correlación; ello, con aplicaciones para la detección temprana de enfermedades, el diagnóstico de una patología, y la clasificación de enfermedades [79] así como subtipos e incluso mutaciones genéticas o virales que pueden representarse mediante análisis secuenciales y aplicaciones de Genómica que potencien la interoperabilidad de los sistemas médicos, el minado de los datos y su obtención mediante aplicaciones del Internet del Todo [73] como parte de un enfoque integral científico y tecnológico.

### 2.2.3. Bioseñales Eléctricas: Electroencefalograma

Dado el enfoque cuantitativo del proyecto de investigación, se optó por emplear el tercer modelo de clasificación [71], donde destacan las bioseñales eléctricas que engloban la medición de actividad cerebral con base a la actividad monitoreada en las neuronas mediante el uso del electroencefalograma [94]:

- **Bioseñales Eléctricas.** También denominadas señales bioeléctricas, cuentan con propiedades y características asociadas a la actividad presente en neuronas y músculos, así como contracciones y fuerza aplicada en estos últimos; ello, de forma no invasiva, puesto que las mediciones se realizan sobre alguna superficie externa del cuerpo humano.
  - **Electrocardiograma (ECG).** Permite monitorear la actividad cardíaca mediante el registro de actividad eléctrica asociada a las funciones que realiza en cuestión de ritmo, intervalos y potencial; cuenta con una frecuencia base de onda de 0.05 a 10 Hz, que puede presentarse en un rango alto de 100 a 1000 Hz.
  - **Electroencefalograma (EEG).** Monitorea la actividad cerebral de una persona, ya sea espontánea o evocada mediante estímulo externo como potencial; posee una frecuencia base de onda de 0.5 a 60 Hz, con subdivisiones en rango de frecuencia *delta* (0.5-4 Hz), *theta* (4-8 Hz), *alfa* (8-13 Hz), y *beta* (13-22 Hz),
  - **Electromiografía (SEMG).** Monitoreo de la actividad muscular como impulso eléctrico mediante electrodos en contacto con el músculo; su frecuencia base de onda puede variar acorde al músculo del que se toma la lectura y el tipo de electrodos; para músculos esqueléticos su rango abarca de 2 a 500 Hz, mientras que para músculo suave cubre de 0.01 a 1 Hz.

Si bien el Electrocardiograma y la Electromiografía proveen datos que pueden contribuir a diferenciar niveles de actividad biológica ante la presencia o ausencia de algún *trigger* con mediciones en tiempo real o de efecto residual, es en el Electrocardiograma donde puede observarse la actividad cerebral de forma concisa y subsecuentemente, la base de la presente investigación para su representación como parte de una interfaz cerebro-computadora [7] para la adquisición, extracción, procesamiento, modelado y clasificación de bioseñales *EEG* [12] mediante el estudio de sus propiedades, aplicando principios de *IA* en aumentación de datos bioeléctricos [25].

## 2.3. Interacción Humano Computadora e Interfaces

La Interacción Humano Computadora o bien *HCI* es parte fundamental del área de desarrollo de la Ingeniería de Software, la generación y transferencia del conocimiento basado en el usuario, así como implementación de soluciones inteligentes con base al conjunto de teorías, métodos y herramientas existentes [78] para la adaptabilidad de parámetros del usuario en entornos educativos y organizacionales; ello tomando el uso de una interfaz adaptativa como medio de referencia para propiciar la interacción con el usuario y la gestión de la retroalimentación al ser personalizable para actividades cotidianas acorde al entorno.

Lo anterior, focalizado a las Interfaces Cerebro-Computadora o *BCI* por sus siglas en inglés, como medio de control y lectura de bioseñales *EEG* [57] asociadas a la representación de los constructos de ansiedad, estrés y temor, como aplicación de enfoque multidisciplinario dentro del contexto tecnológico de la problemática planteada, potenciando el alcance de la investigación y acotando el rango de soluciones acorde al histórico evaluado, lo que se puede censar, su interpretación y almacenamiento [131] para la identificación de parámetros relevantes y actores como parte del diseño instrumental y su modelado.

### 2.3.1. Fundamentos de Interacción Humano Computadora

Los niveles de interacción, al igual que los lenguajes de programación pueden clasificarse de un bajo a un alto nivel con un margen neutro o intermedio como referencia para el monitoreo de cambios subsecuentes en un entorno de desarrollo, de ello que en *HCI* se otorgue prioridad a los niveles de acceso a la información y tareas por desempeñar como criterio para la evaluación de actores involucrados así como procesos correspondientes al acceso y transferencia de datos, flujo de trabajo, dependencias y asociaciones dentro del ambiente digital, las características estáticas y adaptativas de la interfaz, el uso objetivo de la herramienta, e interacciones del mundo real censadas a través de dispositivos móviles. Uno de los aspectos sobre los que parte la interacción y sus métricas, yace en los estudios de la cognición humana como soporte para el diseño lógico y arquitectónico de herramientas basadas en el usuario, aunado al estudio de su entorno, gestión de actividades y requerimientos por cumplir [70].

El aditamento de *HCI* radica en cubrir las brechas existentes entre el desarrollo científico y tecnológico así como entre las dependencias y relaciones de los usuarios, propiciado un acceso universal a recursos y herramientas disponibles como medio de interacción independiente incluso en contraste a la implementación de herramientas asistidas que parten de un modelo general, sin considerar a cada usuario como objeto de diseño [65], requiriéndose entonces de implementaciones basadas en entornos colaborativos para el intercambio de in-

formación y monitoreo de la interacción entre grupos de usuarios así como el análisis de su experiencia compartida e individual del uso de la interfaz, para la aplicación de técnicas de aprendizaje automático que fomenten la adaptabilidad como criterio de diseño e implementación de funciones [121].

Las técnicas mencionadas y el enfoque provisto por *HCI* forman parte del referente sobre el cual, con el objetivo de analizar la validez de intervenciones en áreas con enfoques cualitativos o híbridos como lo es la salud mental, se sustenten revisiones literarias orientadas al ámbito de desarrollo de *software* como parte de la implementación de soluciones asociadas a la adaptabilidad mediante técnicas de aprendizaje automático dada la optimización que proveen a procesos y soluciones existentes así como factibilidad de adopción sistemática [133]; cabe destacar, la contribución a las Ciencias de la Salud mediante la aplicación del enfoque matemático y predictivo en las técnicas de IA y Ciencia de Datos, particularmente en Biomedicina y las vertientes de Bioinformática y Bioestadística, así como en el área de Genómica [89], partiendo de los principios de diseño y experiencia del usuario.

### **2.3.2. Interfaces Cerebro-Computadora**

Los fundamentos de una *BCI* como medio de control para la lectura de bioseñales *EEG* [57] yacen en el estudio fisiológico de las funciones que desempeña el Sistema Nervioso Central (SNC) con el objetivo de trasladar los principios sobre los que operan a un sistema, formando así una *BCI* que permita modificar la interacción del usuario con su entorno [142]; sistemas que si bien se basan en el procesamiento de las señales y las técnicas de filtrado para el conjunto de datos que proveen, han ido adaptándose a la implementación de técnicas algorítmicas de IA orientadas al aprendizaje automático con base a la caracterización del conjunto de datos, decodificación y criterios de selección [67], que pueden expandirse a la implementación de algoritmos bioinspirados, o aprendizaje profundo acorde a la complejidad del caso de estudio así como métricas objetivo [138]. Por otra parte, la Neurociencia coadyuva a la caracterización de funciones desde el punto de vista biológico y cognitivo, englobando el estado de la psique del usuario de la *BCI* y posibles limitantes de adaptabilidad, generando descriptores de la actividad que pueden optimizarse para aumentar la precisión y caracterización de funciones localizadas de las áreas del cerebro humano con enfoques basados en IA y Ciencia de Datos para regiones remotas del cerebro y el estudio de interacción entre diversas áreas como un sistema interconectado de redes [55].

Ha de agregarse al enfoque provisto por las áreas detalladas, la caracterización particular de las bioseñales con las que se desenvuelve el objetivo de la implementación de la *BCI* y se

censa el nivel de interacción del usuario, debido a su carácter no invasivo y la importancia del uso de bioseñales *EEG* en el estudio de procesos sensoriales, cognitivos y motores [111]; por tanto, partiendo de tres etapas primordiales:

- **Obtención.** Toma de datos de bioseñales *EEG* con base a la configuración del dispositivo para su lectura, las regiones del cerebro a monitorear, actividades e instrumentación adicional.
- **Medición.** Lectura de las bioseñales *EEG* con base a la configuración provista de los electrodos para medir los niveles de actividad eléctrica y la respuesta del usuario ante el estímulo.
- **Análisis.** Caracterización de las bioseñales *EEG* mediante un procesamiento previo en que pueden incluirse filtros para mejorar la calidad de las señales por identificar y clasificar.

A la par del desarrollo para *BCI*, en la actualidad se cuenta con adaptaciones a su modelo tradicional mediante la incorporación de modelos múltiples del cerebro, enfoques multisensoriales, y señales multimodales que proveen una mejora en la precisión de la clasificación, al igual que interacciones múltiples e independientes [84]; por otra parte, para el enfoque tradicional se cuenta con implementaciones basadas en el estudio fisiológico de lenguaje corporal y rasgos faciales como complemento a la validación de hipótesis de aprendizaje [90]. Por consiguiente, la importancia de los fundamentos de diseño y principios para la adaptabilidad de *BCI*, así como la representación de los datos obtenidos a través de estas para el estudio de bases de datos existentes y generación de biobancos del conocimiento.

### 2.3.3. Representación del Conocimiento

Considerando la lectura de datos biológicos como una base para la transferencia del conocimiento mediante el uso de bases de datos y su implementación para la generación de biobancos, el estudio del entorno del usuario y nivel de interacción con la interfaz, destaca la transición entre relaciones establecidas por niveles de asociación e interacción, a la representación del conocimiento acorde a características y usuarios objetivo [44], con la finalidad de desarrollar e implementar sistemas que permitan su almacenamiento y generación con técnicas de aprendizaje automático; rasgo extensivo que va más allá de los sistemas expertos, y cubre el uso de sistemas computacionales en diversos dominios de aplicación, fungiendo en el sector salud como herramienta de apoyo para la toma de decisiones, administración

de recursos y medio de representación de registros clínicos [117], proporcionando al usuario múltiples facetas del acceso y tratamiento de la información una vez que se han recabado y procesado los datos, abstrayendo las etapas pertinentes a la relación de este y su interacción en la herramienta así como con otros usuarios.

Adicionalmente, los conjuntos de datos generados como parte de la interacción con diversas interfaces coadyuvan a la implementación de técnicas para la optimización de resultados que abarcan las etapas iniciales del proceso de recolección de datos, su transferencia, y representación [60], validando la escalabilidad de los conjuntos de datos y modelos resultantes, al igual que el uso de técnicas y herramientas para la representación del conocimiento acorde al volumen de los datos y su relevancia. De ello, que su rango de aplicación focalice el monitoreo de condiciones iniciales y su cambio variable así como respuesta ante estímulo para la formulación de hipótesis que, mediante el uso de técnicas de IA provean predicciones que permitan adaptar los tiempos de respuesta en el otro extremo de la interacción, o adaptar la solución a los cambios observados para optimizar el resultado, periodo de respuesta, etcétera [60], culminando en el desarrollo exponencial de soluciones con un alto grado de fiabilidad que puedan emplearse como referente para la solución de problemas complejos basados en la transferencia del conocimiento [107] y optimización con base a indicadores del modelo [146], mismos que bajo el *framework* de Ingeniería de Software pueden desarrollarse con base a modelos y metodologías que incluyan la retroalimentación del usuario para futuras mejoras y transiciones entre sus etapas.

## 2.4. Comportamiento Humano: Salud Mental y Afecciones

Las áreas de estudio del comportamiento humano abarcan los enfoques cualitativos de la observación, y cuantitativos de la implementación del método científico para el análisis, siendo parte de un antecedente histórico por datar y adecuar modelos evolutivos y de conducta como parámetros de adaptación, desarrollo, relación, interacción, afectación, entre otros aspectos de interés para áreas de estudio como la Antropología; de ello, que el desarrollo de aplicaciones centradas en el usuario parta de la integración de equipos multidisciplinarios para cubrir un espectro de aplicación más amplio, en que puedan llevarse a cabo soluciones acordes a múltiples parámetros, siendo un ejemplo la etapa de desarrollo del ser humano, perceptible en la contribución de la Pedagogía a herramientas educativas, empleando técnicas de aprendizaje acorde a los requerimientos del modelo educativo y las necesidades del usuario, adaptando el medio como es el caso de *Learning Paths* en medios virtuales [33].

Partiendo de la vertiente de la medicina y asociado al bienestar de una persona, se puede

introducir el concepto de salud mental como un conjunto de elementos que constituyen el bienestar emocional y psicosocial de un individuo de forma que pueda desenvolverse inter e intrapersonalmente en el entorno que se encuentra, poseyendo herramientas para sobrellevar estrés y factores de riesgo que pudieran comprometer la forma en que se desempeña, siendo extensible a áreas de estudio como la Psicología y Psiquiatría en intervenciones que pueden suscitarse ante la prevalencia de un estado de angustia o sufrimiento, o formar parte de una condición como los son los trastornos mentales o discapacidades psicosociales [105], abarcado los rubros de prevención, análisis, diagnóstico, canalización, tratamiento, seguimiento y reinserción en materia de afecciones mentales. Cabe señalar que uno de los principales factores asociados a la ansiedad es el estrés, sin embargo, también lo es el temor, inclusive la soledad asociada a los efectos del aislamiento y su posible correlación con los ataques de pánico propuestos como caso de estudio; en dicho análisis, un estudio plantea la correlación entre soledad y sabiduría de forma análoga a la capacidad cognitiva y neuronal de un individuo para el manejo de emociones [58]; en las secciones de este capítulo se describe a detalle el efecto de los modelos psicológicos en la conducta humana, así como la caracterización de los trastornos mentales y en específico, de los trastornos del estado de ánimo, la representación de la ansiedad y sus factores clave.

#### **2.4.1. La Conducta Humana e Influencia de los Modelos Psicológicos**

La conducta humana es un constructo que data desde el origen de la humanidad, englobando una forma de actuar, canalizar, o expresar sentimientos que puede ser definida en más de un marco referencial, como lo es el antropológico al hablarse de la humanidad como seres sociables y la forma en que cada ser humano se desenvuelve en sociedad dados factores internos y externos; motivo de estudio de la Psicología al hacerse énfasis en el estímulo que los desencadena, y por consiguiente las implicaciones en Neurociencia Cognitiva para el diseño de experimentos que tienen como objeto el estudio del comportamiento humano a través de bioseñales y el análisis de bibliométricas generadas [124], dando seguimiento al enfoque metodológico y estudio de modelos, incluyendo la referente psicológica en el estudio fisiológico. Preservándose la interrogante del porqué de la conducta en sí, ya que de identificarse el desencadenante puede crearse un modelo para predecir el comportamiento que tendrá una persona ante determinado estímulo, con lo que puede caracterizarse el comportamiento en indicadores psicológicos para detección temprana de anomalías en la conducta de niños y adolescentes [81], fomentando el uso de datos para experimentos multidisciplinarios.

Por otra parte, la respuesta de temor ante una situación apremiante se encuentra basado

en el comportamiento de un individuo, por lo cual entender la forma en que este se conduce resulta de gran importancia para identificar los estímulos que tuvieron parte en ello y modelar los datos obtenidos, siendo escalable a la representación de regiones cerebrales para visualizar los cambios, dando pauta al estudio del consciente y subconsciente.

Neurociencia, Psiquiatría y Psicología son áreas con un enfoque esencial para el estudio del cerebro humano y las afectaciones de las que puede padecer; en Psicología, se cuenta con un antecedente histórico de modelos teóricos que han contribuido a la identificación de rasgos asociados a trastornos mentales y terapia más adecuada acorde a ello. Uno de los principales personajes de la Psicología es Sigmund Freud también conocido como el padre del psicoanálisis, en una de las primeras formas de terapia basándose en las etapas tempranas de vida del ser humano con enfoque a la introspección y percepción de un evento traumático, motivo por el cual su trabajo referencia eventos traumáticos y trastornos asociados [41] o el ámbito interpersonal e intrapersonal como parte del comportamiento [40]; sin embargo, con el paso del tiempo y los avances en el área, dicho modelo fue reemplazado por nuevos enfoques como la Terapia Cognitivo Conductual (*TCC*) que continúa vigente y abarca una aplicación remarcable para casos de ansiedad, extendiéndose a trastornos psicóticos, trastorno bipolar y depresión entre otros, tomando en cuenta el grado de afinidad en comparación a las alternativas [110], como ejemplo de su aplicación se encuentra el trastorno de estrés postraumático, fundamentado mediante aproximaciones teóricas y los mecanismos propuestos con base al comportamiento y la cognición para su tratamiento [99]. Otros modelos son, el de Terapia Gestalt enfocada a una regresión al vientre materno y sus aplicaciones en el cuidado de una persona [6], el Humanista en contraposición a ámbitos epistemológicos de la terapia cuyas bases son el racionalismo y positivismo, siendo entonces una alternativa Psicopatológica [36], y el Sistémico-Familiar como modelo de intervención en terapia, abarcando al paciente, su entorno, el núcleo familiar, y la función del terapeuta [8].

#### **2.4.2. Caracterización de Trastornos Mentales del Estado de Ánimo**

Dentro del grupo de afecciones a la mente se encuentran los trastornos mentales, padecimientos temporales o permanentes que afectan la capacidad mental de una persona al ocasionar cambios en su comportamiento, estos se encuentran clasificados en diversos tipos o clústeres y su diagnóstico puede diferir acorde al grupo de edad al que pertenece la persona, ya que como medida de control algunos diagnósticos diferenciales segmentan no solo por duración de los síntomas en semanas sino por el grupo de edad de la persona dada la complejidad inherente en etapas temprana y las características particulares del trastorno, ello a su vez con

el objetivo de garantizar la veracidad de evaluaciones psicométricas y estudios realizados de forma que se pueda lograr una mejor calidad de vida [134]. Una de las guías para el diagnóstico de trastornos mentales con mayor grado de fidelidad es el *DSM-V, Diagnostic and Statistical Manual of Mental Disorders*, 4.<sup>a</sup> edición [18], donde cabe destacar los trastornos emocionales: trastornos de depresión y trastornos de ansiedad, principalmente estos últimos que se clasifican en [139]:

- **Trastorno de Ansiedad Generalizada.** Ansiedad crónica, con preocupación frecuentemente exagerada a manera de tensión con poco o nulo estímulo.
- **Trastorno Obsesivo-Compulsivo.** Pensamientos o comportamientos recurrentes cuya interrupción eleva los niveles de ansiedad de la persona.
- **Trastorno de Pánico.** Episodios inesperados y repetidos de miedo intenso en conjunto con síntomas físicos que pueden confundirse con un ataque cardíaco.
- **Trastorno de Estrés Postraumático.** Desarrollado tras un evento donde se experimentó miedo extremo o bien se sufrieron daños graves.
- **Trastorno de Fobia Social.** Ansiedad abrumadora acerca de situaciones cotidianas con excesiva autoconciencia de estas.

Acorde a la clasificación presentada anteriormente, los trastornos de ansiedad pueden estar asociados al temor, siendo parte de fobias sociales, dando lugar a una subdivisión de los trastornos de ansiedad con base al indicador [136]:

- **Miedo Primario.** Trastorno de Estrés Post Traumático, Fobia Específica, Trastorno de Ansiedad Social, Trastorno de Pánico.
- **Ansiedad Primaria.** Trastorno de Ansiedad Generalizada, Trastorno de Ansiedad Médica, Trastorno de Ansiedad por Sustancias, Trastorno de Ansiedad por Separación.
- **Otro.** Trastorno Obsesivo-Compulsivo.

### 2.4.3. Representación de la Ansiedad y sus Factores Clave

Como se mencionó en la sección anterior, el DSM-V [18] es una de las herramientas principales para el diagnóstico de trastornos mentales, cubre la clasificación de trastornos, su sintomatología, criterios de diagnóstico diferencial, herramientas adicionales para especialistas de la salud mental, y actualizaciones periódicas realizadas por la Asociación Americana

de Psiquiatría *APA*; no obstante, se cuenta con una equivalencia de los trastornos mentales en la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud *CIE-10* como referente de la Organización Mundial de la Salud, *OMS* [106]. Dado que los trastornos de ansiedad son una de las condiciones con mayor prevalencia, sus métodos diagnósticos se encuentran en constante revisión para brindar el mejor tratamiento farmacológico, cognitivo y de comportamiento, puesto que se encuentran relacionados con el decremento en productividad, aumento en morbilidad y mortalidad, consumo de alcohol y abuso de drogas. Se expone el modelo de ansiedad como [29]:

- **Dominio de los Síntomas.** Relación bilateral entre el sistema de creencias y la señal de alarma, así como con los mecanismos de supervivencia.
- **Proceso.** El rol de la percepción desde el foco de atención y estímulos internos y externos, en relación bilateral con los procesos cognitivos y emocionales con base a la evaluación y memoria, que conllevan a la toma de decisiones como proceso cognitivo y posteriormente a comportamientos como factor emocional.
- **Substrato Neural.** Conexión bilateral entre los ganglios basales y otros circuitos subcorticales con la amígdala y los circuitos de percepción, así como con el córtex y la relación entre estos dos últimos.

Una de las problemáticas que se suscitan al momento de efectuar un análisis preliminar o diagnóstico diferencial es la concepción inicial de los constructos de ansiedad, estrés, y temor, donde la ansiedad puede definirse como un trastorno o estado de ánimo temporal, sin embargo, con frecuencia tiende a asociarse a la presencia de estrés en la persona llegando a generar confusiones con la conceptualización en sí, suceso que se ve ligado al temor; por lo anterior, a continuación en la Tabla 2.1 se muestra una comparación entre ellos para diferenciar entre ansiedad y estrés así como ansiedad y temor [140, 62]:

Tabla 2.1: Caracterización de los Constructos

Características	Constructos		
	Ansiedad	Estrés	Temor
Semejanzas	1. Son respuestas emocionales.		
	2. Pueden presentar insomnio, dificultad para concentrarse, fatiga, tensión e irritabilidad.		
	3. Pueden acelerar los latidos del corazón y la respiración en respuesta al estímulo percibido.		
	4. La actividad física, buena alimentación y cuidado personal coadyuvan a sobrellevar la ansiedad y estrés.		
Diferencias	1. Se desencadena por un estímulo interno.	1. Se desencadena por un estímulo externo.	1. Se desencadena ante peligro en el entorno.
	2. Puede prevalecer inclusive sin estímulo.	2. El estímulo puede ser temporal.	2. El temor moviliza al cuerpo, la ansiedad lo suprime.

De lo anterior se tiene que incluso cuando el estrés y el temor pueden llegar a relacionarse con la ansiedad, existen factores que los diferencian y para el presente proyecto de investigación uno de los principales elementos a aislar es el estrés, ya que el foco del proyecto radica en los ataques de pánico, teniéndose como factores clave la ansiedad que es el padecimiento del que derivan y el temor que es el desencadenante, relación inherente corroborada por lo expresado en el primer y tercer punto de las semejanzas debido a que ambos son respuestas emocionales y los síntomas físicos en común coinciden con los síntomas de un ataque de pánico, añadiendo la sensación de náusea y escalofrío del temor, entre otros.

## 2.5. Metodología de Ciencia de Datos: Aplicaciones de IA

Dando continuidad a la sección anterior, estadísticamente 284 millones de personas padecen de trastornos de ansiedad, de las cuales el 4.7% son mujeres [39], adicionalmente durante la pandemia por *COVID-19* los trastornos de ansiedad en niños y adolescentes mostraron un incremento [66]. El enfoque holístico provisto por el modelo de Ciencia de Datos basado en técnicas algorítmicas para clasificación y categorización de datos es parte clave de procesos para el desarrollo de tecnologías basadas en el usuario, que puedan proveer hallazgos mediante bioseñales *EEG* y la actividad entre regiones cerebrales; la actividad cerebral puede ser difícil de monitorear debido al suceso de un ataque de pánico o incremento de ansiedad puede presentarse de forma imprevisible, puesto que el reflejo fisiológico del cuerpo humano es afectado por estímulos internos y externos [49, 145], por lo que un modelo generalizable de correlación entre eventos puede ser la base para diseños experimentales y casos de estudio.

El diseño experimental propuesto se basa en aplicación de bioseñales *EEG* de pacientes con ansiedad, considerando el tratamiento, historia clínica, diagnóstico de trastorno de ansiedad, y ataques de pánico, para determinar el factor de correlación entre la actividad cerebral y las técnicas estadísticas orientadas a la Ciencia de Datos para el monitoreo de los constructos de ansiedad y su efecto.

### 2.5.1. Metodología en el Marco de Ingeniería de Software

La Ingeniería de Software es una parte elemental de las Ciencias Computacionales en virtud de que moldea soluciones en instancias documentables y esquemas que trasladan prototipos a un ambiente funcional y monitoreable con las interacciones de diversos tipos de usuarios para una evaluación precisa de la mejora continua del ciclo de desarrollo de *software*, asimismo, la implementación de metodologías de *software* permite el monitoreo, optimización y regulación de procesos, motivo por el cual independiente al área de aplicación las metodologías se encuentran adaptadas para satisfacer el análisis de necesidades y requerimientos establecidos.

En la actualidad las metodologías *Lean* y *Agile* lideran en el diseño, desarrollo, e implementación de *software*; sin embargo, la amalgama entre modelos iterativos tradicionales y el foco de diseño presente en la metodología de *Lean UX* [56] provee un medio confiable para la obtención de productos de calidad de manera oportuna y asequible [1] a través de la obtención de *MVPs* como objetivo tras cada iteración a la propuesta inicial, que se adapta con base al *feedback* obtenido. Las aplicaciones metodológicas se extienden en áreas fuera de la industria de *software*, como estudios empíricos que referencian el comportamiento y las experiencias del usuario [76], cubriendo los pilares de *UI* y *UX* facilitando la escalabilidad del *software*, además de aplicaciones en la academia para cubrir recursos, tiempo y costos así como el crecimiento del sector y apremiantes demandas.

El núcleo de la metodología de *Lean UX* se compone de etapas como un todo que forman parte de un ambiente colaborativo para que los miembros del equipo se desenvuelvan [50]; el propósito correspondiente para cada una de sus etapas consta de: el enfoque hacia el problema que establece un marco común para la estrategia a seguir, el curso de acción propuesto para el enfoque hacia el problema que ayude a contextualizar los pasos a tomar, el diseño de los *MVPs* con un potencial de escalabilidad, la ejecución que puede ser vista como un diseño experimental, el caso de estudio o prueba acorde a la iteración en curso y el grado de avance, y el *feedback* que impulsa los avances hacia la meta objetivo a la par que propicia la transición a otra iteración del ciclo de así requerirse.

### 2.5.2. Ciencia de Datos para la Resolución de Problemas

Las ramas de IA y Ciencia de Datos proveen la base para el desarrollo de tecnologías modernas y soluciones basadas en datos [4, 123], siendo que la Ciencia de Datos abarca a la IA y *Big Data* entre otras áreas, mientras que la IA cumple con propósitos específicos dentro del nicho; dado el énfasis en la clasificación y categorización de los datos así como a etapas posteriores al análisis inicial, el enfoque para el modelo de correlación se basó en una implementación temprana de técnicas algorítmicas de IA para la clasificación y categorización del conjunto de datos, con un desarrollo basado en los principios de la Ciencia de Datos para la exploración, limpieza, análisis, procesamiento, y perfilado de estos.

Al ser la base sobre la cual la programación y desarrollo de aplicaciones pueden ser llevados a un alto nivel de implementación dada su complejidad para proveer una solución acorde a la problemática, las técnicas algorítmicas deben ser investigadas, evaluadas, y comparadas para seleccionar aquella acorde al conjunto de requerimientos, considerando a su vez sus principios matemáticos y la posibilidad de incurrir en causalidad dada la inferencia [109]; el modelo de correlación propuesto fue construido con el tiempo y uso de recursos como prioridades principales, lo que condujo a la complejidad de desarrollo e implementación de las técnicas algorítmicas a ser considerados como parte de los requerimientos de selección, subsecuentemente pese a la capacidad de las redes neuronales para trabajar con grandes cantidades de datos y su afinidad a la problemática dada la flexibilidad de su *kernel* [63], al considerar los requerimientos del diseño metodológico e instrumental para el desarrollo del modelo e implementación algorítmica, fue descartado su uso para la etapa de desarrollo actual. Dando seguimiento al proceso de selección [24, 9], en adhesión a los requerimientos y orientación a la solución de problemas, los principales algoritmos de IA para clasificación y categorización fueron estudiados con el objetivo de determinar cuáles usar para la exploración, obtención de hallazgos, y modelado del conjunto de datos; la Tabla 2.2 detalla las técnicas algorítmicas que fueron analizadas para la problemática actual; denotando el modelo al que pertenecen, una breve descripción de estas, y comentarios del proceso de selección para el modelo de aprendizaje supervisado, y de aprendizaje no supervisado:

Tabla 2.2: Ciencia de Datos: Técnicas Algorítmicas de IA

<b>Aprendizaje</b>	<b>Modelo</b>	<b>Descripción</b>	<b>Comentarios</b>
Supervisado	<i>Decision Tree</i>	Árbol de decisiones basado en características.	Uso de valores nulos y sensibilidad a valores atípicos.
Supervisado	<i>Random Forest</i>	Conjunción de árboles y sus múltiples resultados.	Alta precisión y complejidad de entrenamiento.
Supervisado	<i>Gradient Boosting Regression</i>	Predicciones con base al aprendizaje más débil.	Alta complejidad y costo computacional.
No Supervisado	<i>K-Means</i>	Agrupamiento por distancia Euclidiana.	Agrupaciones reducidas dado un número de $K$ .
No Supervisado	<i>Hierarchical Clustering</i>	Agrupamiento con un enfoque ascendente.	Alta complejidad que no selecciona el mejor clúster.
No Supervisado	<i>Gaussian Mixture Models</i>	Agrupamiento para datos normalmente distribuidos.	Afinamiento complejo requerido para óptimos resultados.

### 2.5.3. Técnicas Algorítmicas: Clasificación y Categorización

En la actualidad se cuenta con una amplia colección de técnicas algorítmicas dado el influjo de información y clasificación que permiten su mejora, ya sea como una optimización mediante la creación de otro algoritmo basado en una técnica específica, la combinación de técnicas o la propuesta de nuevas; sin embargo, adicional a los requerimientos de implementación por cubrir y los criterios de selección acorde al problema, es importante identificar de antemano el área de implementación y características intrínsecas. Por lo tanto, aunque en la Tabla 2.2 de la sección anterior se muestra que la técnica de *Random Forest* provee una mayor precisión, también incrementa la complejidad de su entrenamiento comparado a la técnica de *Decision Tree*, particularmente *ID3* y *J48* respecto al enfoque supervisado, mientras

que para el aprendizaje no supervisado, para organización de datos la técnica de *Hierarchical Clustering* no siempre selecciona el mejor clúster, más sí incrementa la complejidad a comparación de la técnica de *K-Means* que puede implementarse con un determinado número de *K* a escalar sin aumentar la complejidad de su interpretación.

### Clasificación de Datos: Algoritmo *K-Means*

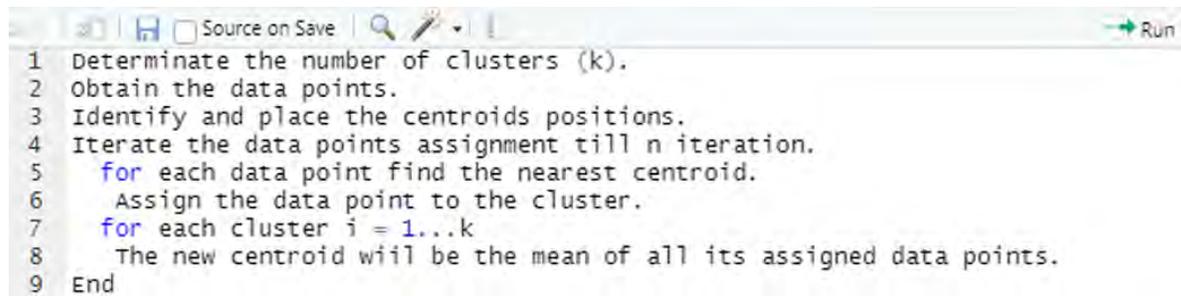
La identificación de patrones puede visualizarse acorde a la distribución de los puntos de datos mediante un modelo de regresión lineal, pero sin procesamiento adicional los grupos correspondientes y sus características solo pueden ser teorizados, motivo por el cual la clasificación de los datos es requerida; *K-Means* es un algoritmo de agrupamiento basado en el etiquetado de datos mediante la determinación del *K* número de clústeres a emplear y el proceso iterativo para identificar sus centroides, posicionando los puntos de datos en el clúster correspondiente [27]. Su implementación da solución a un problema de optimización bajo la función de minimizar, y opera aplicando el método general propuesto por Wong, H. en 1979 [61], que define al total de la variación intra-clúster como la suma de distancias al cuadrado por elemento, respecto a su centroide, tal y como se muestra en la Ecuación 2.1:

$$W(C_k) = \sum_{x_i \in C_k} (x_i - \mu_k)^2 \quad (2.1)$$

Para cumplir con la función objetivo de minimizar, cada elemento es asignado a un clúster de forma tal que la suma al cuadrado de las distancias de los elementos a su centroide sea minimizada, permitiendo la obtención del total de la variación de las distancias al aplicar la Ecuación 2.2:

$$\sum_{k=1}^k W(C_k) = \sum_{k=1}^k \sum_{x_i \in C_k} (x_i - \mu_k)^2 \quad (2.2)$$

Consecuentemente, evaluando la tendencia del conjunto de datos respecto a la varianza obtenida; de forma puntual, la Ecuación 2.1 y la Ecuación 2.2 son parte del proceso interno de ejecución del algoritmo, cuyo planteamiento se muestra como pseudocódigo en la Figura 2.2:



```

1  Determinate the number of clusters (k).
2  Obtain the data points.
3  Identify and place the centroids positions.
4  Iterate the data points assignment till n iteration.
5  for each data point find the nearest centroid.
6  Assign the data point to the cluster.
7  for each cluster i = 1...k
8  The new centroid will be the mean of all its assigned data points.
9  End

```

Figura 2.2: Pseudocódigo para el Algoritmo de *K-Means*.

Por tanto, para su implementación es necesario definir en primera instancia el número de clústeres  $K$  que se requiere para el conjunto de datos en cuestión, y el número de centroides que contienen; con  $K$  generado mediante la aplicación del Método de Codo, posterior a lo cual para conocer el número de centroides en cada clúster se aplica la fórmula de la Distancia Euclidiana, acorde a la distribución de los puntos de datos y sus tendencias [86, 23]. El Método de Codo consiste en la asignación incremental de valores a  $K$ , dando como resultado el número óptimo de clústeres para *K-Means* bajo el preámbulo de que conforme incrementa el valor de  $K$  la distorsión presente en el conjunto de datos disminuirá y se acercará a un comportamiento promedio, cuyo punto de inflexión denominado por la forma observada en el gráfico que genera denota el número óptimo para  $K$  [38]; puntualizando la minimización de la variación intra-clúster mediante su suma de cuadrados, denotada por sus siglas en inglés como WSS, la cual es obtenida mediante la Ecuación 2.3:

$$\text{minimizar} \left( \sum_{k=1}^k W(C_k) \right) \quad (2.3)$$

Por otra parte, el cálculo de la distancia para el conjunto de datos puede variar acorde a ello, pudiendo implementarse la Distancia de Manhattan o la Distancia de Canberra [47]; sin embargo, el uso de la Distancia Euclidiana predomina sobre el de estas, tal y como es el caso para el presente trabajo de investigación; la Distancia Euclidiana se representa acorde a la Ecuación 2.4 [14]:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - y_{jk})^2} \quad (2.4)$$

En la que la distancia  $d$  entre dos puntos  $i, j$  con un centroide  $i$  y atributo  $j$  para datos representados por  $K$  proporciona la distancia media entre los puntos y sus centroides, moldeando los clústeres que se formaran del conjunto de datos con cada iteración.

El enfoque provisto por el algoritmo de *K-Means* es clave para la adaptación hacia una aplicación no supervisada con *Random Forest* o *Bagging*, inclusive su implementación en conjunto con técnicas como redes neuronales [19]; particularmente, el modelo propuesto se enfoca en el Método de Codo para validar el número de centroides y las tendencias de datos observados para que puedan ser utilizadas como variables de control del conjunto de datos y sean un punto de comparación respecto a su categorización.

### Categorización de Datos: Algoritmos *ID3* y *J48*

La toma de decisiones puede beneficiarse de clases predefinidas o datos categóricos en un conjunto de datos, por ende la técnica de *Decision Trees* requiere de al menos un tipo de dato categórico para generar predicciones basadas en las variables correspondientes por categoría; aunque *Random Forest* es una de las técnicas principales aplicadas cuando un árbol de decisión es requerido, los principios clave de la categorización mediante la creación de nodos y hojas así como la jerarquía de padres e hijos son optimizados con la Ganancia de Entropía para definir cuál variable tiene el peso más alto, de forma que pueda ser asociada a la categoría con un rango proporcional de precisión respecto a su pertenencia y parámetros a considerar [130]. El algoritmo *ID3*, propuesto por *Quinlan, R.* en 1986, implementa un árbol de decisión *m-way* basado en dicho preámbulo y el enfoque *greedy* para encontrar el atributo de categorización con la mayor Ganancia de Entropía [112]; requiriendo de conjunto de ejemplos conformado por subconjuntos positivos y negativos, como se observa en la Ecuación 2.5:

$$I(p, n) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n} \quad (2.5)$$

Los cuales a su vez constituyen el promedio de los pesos del árbol a conformar con base al atributo que funge como nodo raíz *A*, mediante la Ecuación 2.6:

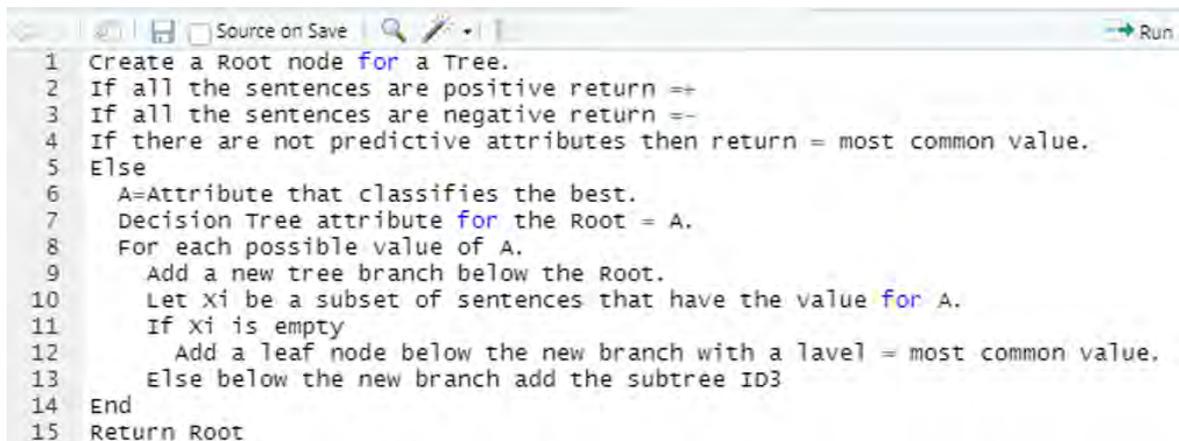
$$E(A) = \sum_{i=1}^v \frac{p_i + n_i}{p+n} I(p_i, n_i) \quad (2.6)$$

Que provee una segmentación concisa de las relaciones, en que la proporción de pesos de sus ramas representa la Ganancia esperada de la Entropía, denotada en la Ecuación 2.7:

$$\text{Ganancia}(A) = I(p, n) - E(A) \quad (2.7)$$

Por consiguiente, parte clave del árbol de decisión *ID3* consta del conocimiento y aplicación de las métricas de Entropía e inherente Ganancia; la Figura 2.3 muestra a detalle el

Pseudocódigo a seguir para su implementación:



```

1 Create a Root node for a Tree.
2 If all the sentences are positive return ==
3 If all the sentences are negative return ==
4 If there are not predictive attributes then return = most common value.
5 Else
6   A=Attribute that classifies the best.
7   Decision Tree attribute for the Root = A.
8   For each possible value of A.
9     Add a new tree branch below the Root.
10    Let xi be a subset of sentences that have the value for A.
11    If xi is empty
12      Add a leaf node below the new branch with a label = most common value.
13    Else below the new branch add the subtree ID3
14 End
15 Return Root

```

Figura 2.3: Pseudocódigo para el Algoritmo *ID3*.

Retomando los fundamentos de las métricas, particularmente de la Entropía, término acuñado por *Shannon, C.* [120], aunado a la existencia de valores no disponibles, u atributos continuos en el rango de datos por analizar, la estrategia para la resolución de problemas de *ID3* puede no ser la óptima; aunque, sus resultados pueden ser mejorados al implementar técnicas de optimización [144] que conduzcan a una implementación con reducción significativa de su margen de error el cual es de gran importancia dado el contexto de datos médicos involucrados en el modelo propuesto. Con base a la derivación de reglas y la aplicación de *prunning* para evitar el riesgo de *overfitting* en determinados escenarios, *Quinlan, R.* propuso una versión optimizada de *ID3*, el algoritmo *C4.5* [113] con las características de su predecesor, y un énfasis en la noción adquirida de los atributos a través de su Ganancia, reduciendo el uso de recursos computacionales al generar árboles de decisiones más pequeños para identificar atributos, ejecutando los procesos de asignación y ordenamiento acorde al nivel de Ganancia obtenida [85]; favoreciendo así no solo atributos que posean un número considerable de valores, sino balanceando los registros por atributo al mediante la proporción de su Ganancia, como se muestra en la Ecuación 2.8 :

$$Proporcion_i = \frac{p_i + n_i}{\sum_i (p_i + n_i)} \quad (2.8)$$

Donde la distribución de los elementos se encuentra dada por los valores asignados al atributo raíz bajo la función de optimización de minimizar, clasificando la información con la Ecuación 2.9:

$$IV(A) = - \sum_{i=1}^v \frac{p_i + n_i}{p + n} \log_2 \frac{p_i + n_i}{p + n} \quad (2.9)$$

Con lo que puede procederse a la evaluación de puntos de inflexión desde las ramas generadas en el proceso y sus nodos, hasta la estructura final; por otra parte, el enfoque particular del algoritmo puede ser ajustado acorde al *framework* de desarrollo y herramientas a emplear, así como el tipo de licencia y derechos de autor, motivo por el cual para el proyecto de investigación se optó por emplear la herramienta de Weka e implementar la versión *Open Source* del *C4.5*, el algoritmo *J48* [125] que denota la mejora en la adquisición de información, y métrica de Entropía como factor diferencial para corroborar los resultados obtenidos con el algoritmo de *K-Means* dado su alto rango de precisión, la posible identificación de factores de riesgo y su representación concisa en la matriz de correlación que puede generar [122].

#### 2.5.4. Comparación de Metodologías y Caracterización

En Ingeniería de Software, una de las funciones principales de las directrices yace en identificar a los actores principales de una problemática, dando seguimiento a interacciones para determinar el flujo de trabajo de los requerimientos por evaluar en la solución objetivo, siendo factible un seguimiento metodológico para cubrir las necesidades del usuario mediante un enfoque adaptable que replique el modelo de la situación, productos obtenidos y soluciones. Sin embargo, no todas las metodologías presentan una solución óptima, incluso si proveen los resultados objetivo, por lo que el enfoque de áreas involucradas deben ser evaluadas minuciosamente para evitar incurrir en un procedimiento obsoleto o una filosofía inadecuada que pueda comprometer el tiempo invertido en el proyecto, tareas por completar, recursos a emplear, instrumentos a aplicar, y el capital humano y financiero.

El enfoque propuesto para esta investigación engloba la metodología de *Lean UX* [56] y la metodología de Ciencia de Datos de *IBM* [118], unificando el planteamiento analítico de la problemática con los datos de entrada provistos por el usuario, el nivel de sus interacciones y la obtención de resultados parciales mediante la estructuración y segmentación de los datos crudos culminando en la creación de modelos y asignación de procesos, cubriendo las bases teóricas de un producto de *software* basado en el usuario, y la practicidad de los métodos y técnicas computacionales.

### Lean UX Methodology

Las metodologías desempeñan un papel protagónico para el diseño, desarrollo, e implementación de *software* en conjunto con el grado de aceptación de los usuarios y *stakeholders*; sin embargo, cada metodología se estructura acorde al área en que se implementará y los requerimientos por cubrir, como ejemplo de ello los agentes inteligentes, el análisis y regulación de sus datos [53, 2]. Comúnmente, las metodologías *Lean* y *Agile* son aplicadas como un modelo generalizable para brindar solución a una amplia gama de problemas, lo que dio pauta a la combinación de modelos iterativos tradicionales y diseños orientados al usuario, como lo plantea la metodología *Lean UX* [56] que brinda al usuario productos de calidad sin comprometer los requerimientos de costo y tiempo [1] generando *MVPs* por cada iteración, mismos que pueden ser mejorados con el *feedback* recolectado, siendo el proceso clave del ciclo que la fundamenta, como puede observarse en la Figura 2.4:

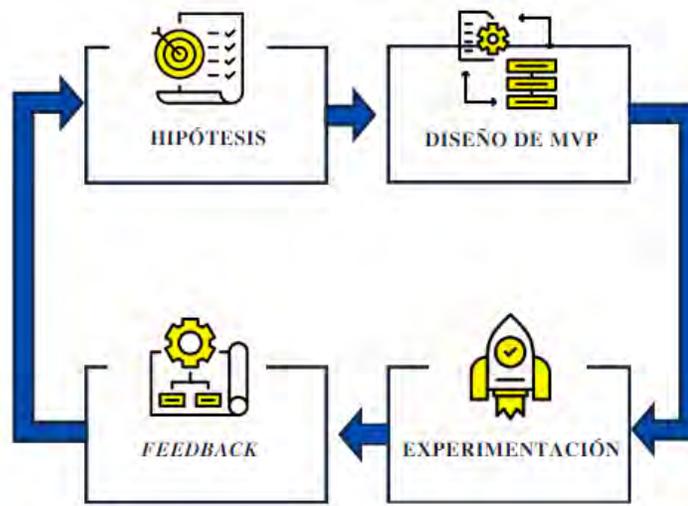


Figura 2.4: Ciclo de la Metodología de *Lean UX* [56].

El rango de sus aplicaciones cubre no solo la industria de *Software*, sino áreas en que la mejora de procesos se sustenta en la optimización de los requerimientos de costo y tiempo para mejorar la experiencia del usuario, considerando su comportamiento [76] e implicaciones de la escalabilidad del producto de *Software* para cubrir las demandas del usuario, el papel de cada miembro del equipo [50], las tareas clave por completar en cada etapa y el enfoque a emplear acorde al diseño experimental, caso de estudio o prueba orientada a avances respecto a la meta objetivo y transiciones por cada ciclo o iteración.

### IBM Data Science Methodology

La Ciencia de Datos es un área derivada de las Ciencias Computacionales que ha crecido remarcablemente dado el incremento constante en el volumen de los datos, lo cual se refleja en sus campos de aplicación, por ende la existencia de metodologías y soluciones basadas en datos para el cumplimiento de objetivos, o bien aplicaciones basadas en la interconexión de dispositivos y el origen de los datos, especificaciones que conlleven a retos para la analítica de datos y aplicación de técnicas [91]; matemáticamente hablando, una de las áreas con alta demanda para metodologías de Ciencia de Datos y su particular implementación, yace en la estadística [17] dada la aplicación de conocimiento existente para la obtención de nuevas perspectivas y creación de indicadores en cada área cubriendo técnicas básicas y complejas. Particularmente, esta investigación se condujo bajo el enfoque de la Metodología de Ciencia de Datos de *IBM* [118] ejemplificada en la Figura 2.5:



Figura 2.5: Flujo de la Metodología de Ciencia de Datos de *IBM* [118].

Cuyo criterio analítico de la problemática prioriza el flujo de datos en etapas tempranas, así como su análisis y culminación en el *feedback* para optimizar la solución alcanzada; el enfoque de Ciencia de Datos hace de la metodología de *Lean UX* [56] un *framework* de referencia para la evaluación, implementación y *feedback* obtenido como parte de un ciclo iterativo basado en datos para la solución de problemas, mediante la conexión entre la interacción del usuario y el flujo de los datos para el caso de estudio de la metodología propuesta.

# TESIS TESIS TESIS TESIS TESIS

## Capítulo 3

### Metodología

Dada la naturaleza de la problemática, previa formulación de la metodología se realizó un análisis del estado del arte en las vertientes involucradas; retomando el área de Salud Mental como foco de aplicación en materia de ataques de pánico, metodologías y aplicaciones orientadas a ello. Independiente al enfoque metodológico y propuesta de modelos para soluciones orientadas al área de la Salud Mental, la perspectiva de las contribuciones puede variar respecto a la representación de la ansiedad, sus indicadores y bases tecnológicas por cubrir.

Tomando a la ansiedad como el indicador clave, debe mencionarse que el estudio de las funciones cognitivas del cerebro humano y capacidad para regular emociones puede proveer datos relevantes para la identificación de estados emocionales y sus efectos en el comportamiento, por lo que los trabajos relacionados en Ciencias de la Salud que se cubrieron como complemento al enfoque de las Ciencias Computacionales, abarcan el paradigma de otros tipos de trastornos de ansiedad, las técnicas implementadas como referente de la elección de algoritmos, así como otras tecnologías y áreas de aplicación relacionadas que pueden ser consideradas para trabajos a futuro. La Tabla 3.1 muestra el desglose de los principales trabajos relacionados que fueron considerados, detallando el sustento tecnológico que fue aplicado a los constructos de ansiedad en conjunción con los procesos cognitivos y emocionales abordados en la descripción de la metodología u aplicación.

Tabla 3.1: Ataques de Pánico: Metodologías y Aplicaciones Relacionadas

Referencia	Tecnología	Constructo	Descripción	Enfoque
Khessiba, S et al. [74]	<i>Deep Learning</i>	Inferencia	Estudio de la actividad cerebral mediante <i>Bioseñales EEG</i> y arquitecturas de <i>Deep Learning</i> .	<i>Neural Computing</i>
Chen, C et al. [35]	<i>Bioseñales EEG</i>	Ansiedad	Evaluación de <i>Neurofeedback</i> para el alivio de Ansiedad.	Aplicaciones de Neuro-psiquiatría
Balan, O et al. [20]	<i>Machine Learning, Deep Learning</i>	Temor	Investigación de técnicas para automatizar la detección de niveles de Temor.	Inteligencia Artificial y Realidad Virtual
Suhaimi, N et al. [129]	<i>Bioseñales EEG</i>	Cognición	Identificación de estados emocionales del ser humano utilizando <i>Bioseñales EEG</i> .	Inteligencia Computacional y Neurociencia
Beaurenaut, M et al. [22]	Cuestionario STAI	Ansiedad	Paradigma <i>Threat of scream</i> para el estudio de Ansiedad fisiológica y subjetiva.	Correlación Estadística
Francese, R et al. [51]	Realidad Virtual	Emociones	Metodología basada en el usuario para un sistema de detección de emociones.	<i>User Experience</i>

Si bien el análisis de la revisión literaria detallado en la Tabla 3.1 destaca aplicaciones tecnológicas y metodológicas para ansiedad, principalmente en técnicas de IA, aplicaciones de Neuropsicología y Neurociencia, es importante considerar el grado de impacto de la metodología en sus etapas iniciales, y su capacidad de crecimiento exponencial. Este capítulo abarca las particularidades de la propuesta metodológica, la guía de diseño de sus *MVPs*, y la estructura del modelo obtenido a partir de esta; dando prioridad al enfoque de *Lean UX* a la par del análisis del flujo de datos mediante la Ciencia de Datos, lo que provee valores de referencia para especialistas de la salud, con base a percepciones obtenidas que pueden derivar en la optimización de procesos e identificación de factores de correlación.

### 3.1. Propuesta Metodológica

Considerando el antecedente de *Lean UX* en conjunto con el enfoque híbrido de *BCI* y la contextualización de ataques de pánico, el desarrollo de la metodología se centró hacia un proceso iterativo concebido para la segmentación de usuarios y cumplimiento de los requerimientos para pruebas, previa recolección de datos para la elaboración de un caso de estudio y obtención de *feedback*; como puede observarse en la Figura 3.1, que introduce las etapas de la metodología en una secuencia lineal: Segmentación, Diseño, Tratamiento, Lanzamiento y Evaluación, con un esquema basado en el desglose dual de subsecciones correlacionadas para cada etapa, cuyos procesos se iteran mediante el vínculo entre la etapa de Segmentación y la etapa de Evaluación.

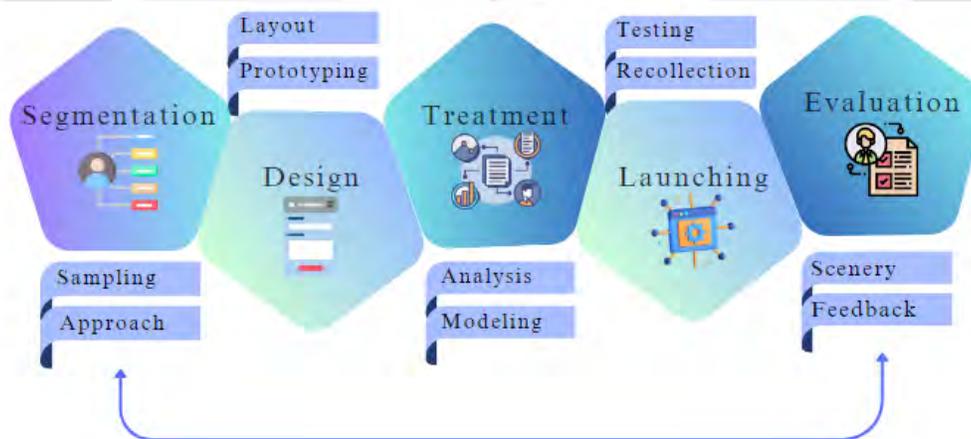


Figura 3.1: Desglose del Diseño Metodológico [31] Inspirado en *Lean UX* [56].

### 3.1.1. Segmentación: Tamaño de la Muestra y Enfoque

La segmentación es la etapa de apertura de la metodología y la base para definir la población, usuarios objetivo, requerimientos subsecuentes y necesidades por cubrir en determinado número de iteraciones; para cumplir con los objetivos proyectados en un margen de tiempo para la revisión y aprobación de los *MVPs*, se encuentra dividida en muestreo y enfoque:

- **Muestreo.** Identificación inicial de la población muestra con base a indicadores estadísticos y demográficos para su posterior clasificación; delimita el esquema de trabajo para el tipo de usuarios, proveyendo una estructura para la asignación de recursos y papeles a desempeñar por el equipo multidisciplinario.
- **Enfoque.** Tipo de investigación y técnicas propuestas para aplicar a la población objetivo dados los indicadores y papeles identificados en el muestreo, define los requerimientos por cubrir en cada una de las etapas subsecuentes e iteraciones.

### 3.1.2. Diseño: Esquematización y Prototipado

La etapa de diseño es el punto central de la metodología y los *MVPs* a ser entregados en la última etapa, incluye la representación de los requerimientos en diseños y esquemas de las tareas principales, así como de sus procesos, dando seguimiento a los avances y flujo de trabajo, conduciéndolos hacia una implementación temprana a nivel de prototipado; se encuentra dividida en esquema y prototipo:

- **Esquema.** Integración de diagramas de Ingeniería de Software para la representación de papeles de los usuarios y su clasificación, del equipo multidisciplinario, del sistema y producto de *software* como un todo. El objetivo del esquema es la identificación del flujo de trabajo de los procesos para su vinculación con los requerimientos, así como optimización en medida de lo posible, sustentando el *framework* de referencia.
- **Prototipado.** Desarrollo inicial de la solución de *software* por implementar en etapas subsecuentes, considerando casos de uso y el progreso para el cumplimiento de los requerimientos, ajustando de forma consistente el nivel de funcionalidad del prototipo. Moldea los avances de los *MVPs* para obtener nuevas perspectivas del equipo de multidisciplinario e incluirlas en los ajustes por efectuar.

### 3.1.3. Tratamiento: Análisis y Modelado

La etapa de tratamiento involucra un análisis exploratorio de los datos recabados del prototipo elaborado en la etapa de diseño, para la obtención y aplicación de un modelo de datos para efectuar pruebas y recabar datos previo lanzamiento, de forma operativa. El propósito de la etapa de tratamiento consiste en el proceso de limpieza efectuado entre las subetapas para monitorear el margen de error y cambios en el conjunto de datos, por lo que se divide en análisis y modelado:

- **Análisis.** Procesos y técnicas orientados a datos para la exploración de un determinado conjunto de datos, identificación de su estructura, composición y tendencias; su objetivo consiste en proveer hallazgos de los datos obtenidos del prototipo, monitoreando avances y detectando áreas de oportunidad.
- **Modelado.** Transformación del conjunto de datos obtenido en la etapa previa para efectuar la vinculación entre los requerimientos con los hallazgos del análisis; su objetivo es perfilar la solución de *software* y aplicar la mejora continua al modelo que se desprende de esta, a la par que asigna instancias y funciones.

### 3.1.4. Lanzamiento: Experimentación y Recolección

La etapa de lanzamiento se enfoca en el desarrollo de la solución de *software* que permite la evaluación de la hipótesis formulada al inicio del proceso, corroborando los objetivos del espacio de investigación propuesto; la solución de *software* debe ser probada por los usuarios, y nuevos datos de la interacción deben ser recolectados y almacenados, implicando su experimentación y recolección:

- **Experimentación.** Prueba inicial de la solución de *software* basada en una *BCI* para la ejecución de pruebas de campo, rendimiento, monitoreo del factor humano y aspectos computacionales; tiene como objetivo brindar un espacio para la observación y análisis que permita la identificación y ajuste de errores que puedan ocurrir, sugiriendo cambios, optimizando áreas, o bien aprobando el lanzamiento del *software*.
- **Recolección.** Recolección de datos con enfoque central en el usuario y el dispositivo que conforman la *BCI*; su propósito consiste en la observación objetiva del comportamiento para la medición del nivel de interacción con el usuario, focalizando los pilares de la metodología de *Lean UX*, *UX* y *UI*.

### 3.1.5. Evaluación: Simulación y *Feedback*

La etapa de evaluación es la culminación del flujo lineal de datos de la metodología a través de los puntos de verificación de sus cinco etapas clave, y el inicio de su proceso iterativo, que engloba los *MVPs* obtenidos en cada una de las etapas al igual que la solución óptima obtenida; la evaluación se conforma por la definición del caso de estudio mediante la simulación y la incorporación del *feedback*:

- **Simulación.** Datos representativos obtenidos de los procesos de experimentación y recolección, en conjunto con la aplicación de indicadores estadísticos y demográficos con base a la población muestra de usuarios objetivo; su propósito es contextualizar y crear una base para el modelo de comportamiento del usuario.
- ***Feedback.*** Datos provistos por los usuarios y miembros del equipo multidisciplinario respecto al desarrollo del *software*, cubriendo la evaluación de los requerimientos y las sugerencias de aplicación para futuras implementaciones; el objetivo del *feedback* consta de la adopción de cambios oportunos que puedan reflejarse en las interacciones del usuario y recursos empleados.

## 3.2. Guía de Diseño de *MVPs*

Como parte de los objetivos formulados en cada una de las etapas bajo las premisas de *Lean UX*, una iteración interna fue ejecutada para los miembros del equipo multidisciplinario involucrados en las etapas iniciales del diseño metodológico: el desarrollador, el especialista de la salud, y el analista de datos; *MVPs* fueron obtenidos mediante el proceso, en conjunto con el *feedback* correspondiente para sustentar las bases del conocimiento de la propuesta y los productos como un preámbulo para la implementación algorítmica y casos de estudio activos de ataques de pánico. Esta sección y las subsecciones que le conforman muestran el diseño analítico de los avances obtenidos, describiendo la investigación realizada para cada una de las etapas de la metodología, los *MVPs* obtenidos, el nivel de logro alcanzado, y notas puntuales para su mejora; adicional, a un filtro entre el *feedback* como un todo para futuras aplicaciones en intervenciones psicológicas en el municipio de Aguascalientes, México en el Sistema para el Desarrollo Integral de la Familia (DIF), entidad gubernamental orientada a brindar apoyo social a la población en áreas de la salud como rehabilitación, apoyo psicológico y cuidado médico [126, 43].

### 3.2.1. Segmentación: Enfoque del Equipo Multidisciplinario

Para validar el análisis de la población muestra y usuario objetivo, el flujo de trabajo del equipo multidisciplinario para el desarrollo del *software* fue detallado acorde al esquema general de los datos y procesos de actividades a realizar; en la Figura 3.2 se muestra como actor al usuario objetivo, cuya interacción con el especialista de la salud permite realizar una evaluación que posteriormente es recibida por el ingeniero *DevOps* que en etapas tempranas de desarrollo caracteriza el conjunto de datos, previo modelado por el científico de datos, que mediante la implementación del modelo obtiene *feedback* del perfil del usuario, quien al igual que los miembros del equipo multidisciplinario, puede proporcionar *feedback*.

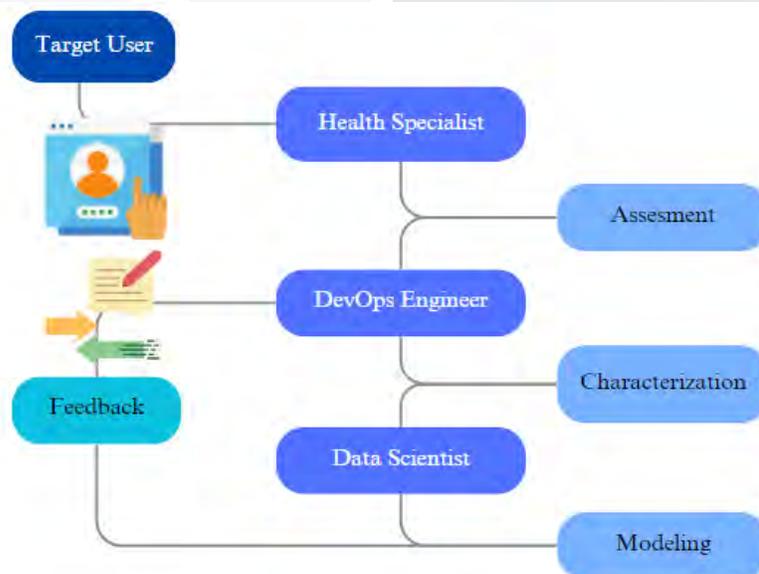


Figura 3.2: Esquema del Equipo Multidisciplinario.

Tras una revisión de antecedentes del perfil psicológico, la población muestra fue identificada con base a la asociación de los indicadores más relevantes: edad, condición de salud, y síntomas; consecuentemente, la población muestra se segmentó en personas de 15 a 40 años con trastorno de ansiedad, ataques de pánico, y/o síntomas relacionados, a ser abordada en un estudio de proceso cuantitativo.

### 3.2.2. Diseño: Esquema del Sistema a Prototipar

Con el objeto de cubrir las interacciones entre los componentes clave involucrados en el desarrollo del *software*, un diseño arquitectónico de diagramas de clase fue elaborado; mostrado

en la Figura 3.3, se encuentra la representación de la interacción entre los dos procesos principales del sistema en un nivel de abstracción de componentes, con clases asignadas para el análisis y procesamiento respectivo de las bioseñales *EEG*, así como los datos de entrada del usuario y los datos de salida para el especialista de la salud.

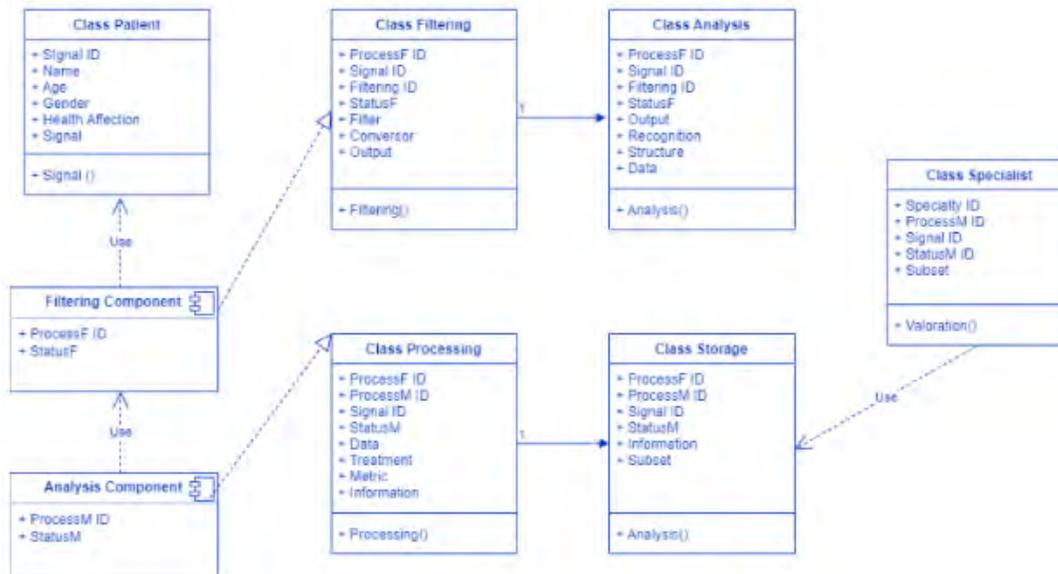


Figura 3.3: Diagrama de Clases del Sistema.

El flujo de datos se origina en la clase de Paciente, puesto que integra los indicadores clave de edad y condición de salud acorde a los criterios de segmentación de la población muestra, aunado al nexo entre los procesos de las bioseñales *EEG* por convertir, estructurar, y tratar, al igual que el subconjunto de datos para el uso del especialista de la salud.

### 3.2.3. Tratamiento: Análisis y Modelado de Datos

Las funciones del sistema fueron delimitadas en un diagrama de contexto tal y como se muestra en la Figura 3.4 que denota la *BCI*, los componentes clave para el desarrollo del *software* y el tipo de usuario, ejemplificando la relación entre los componentes mediante el flujo de datos de entrada y de salida, con el actor principal siendo el Paciente del cual se obtienen las bioseñales *EEG*, así como el especialista de la salud como el usuario final que recibe el acceso a los datos obtenidos de las bioseñales una vez que estas han sido procesadas adecuadamente, filtradas y analizadas; por consiguiente, ofreciendo como resultado datos esenciales que pueden ser almacenados para acceso futuro.

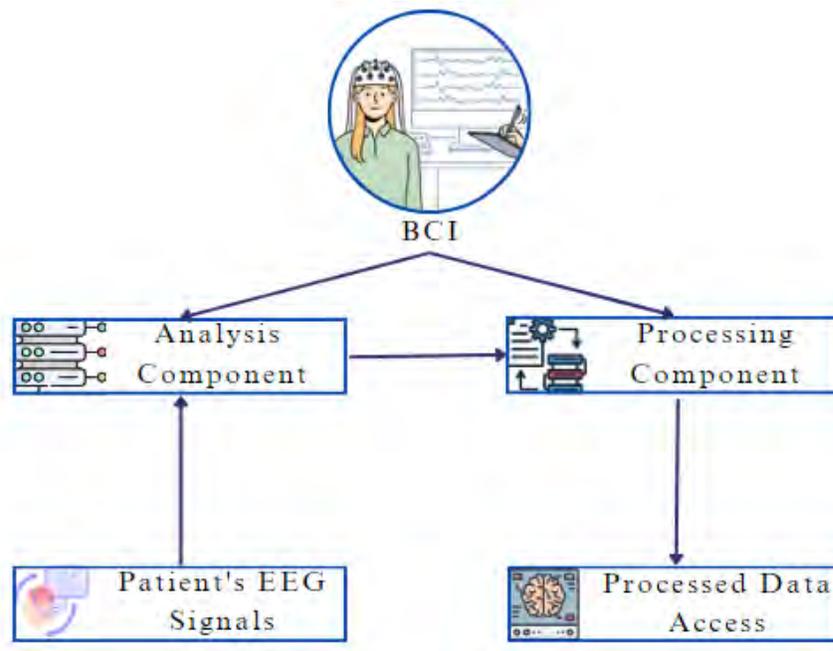


Figura 3.4: Diagrama de Contexto del Sistema.

### 3.2.4. Lanzamiento: Pruebas de uso de la *BCI*

Complementario al Diagrama de Contexto mostrado anteriormente en la Figura 3.4, el Diagrama de Contenedores de la Figura 3.5 demuestra la relación entre los componentes de *software* en un nivel técnico, tomando en consideración la abstracción entre el proceso de tratamiento de los datos donde los datos de entrada de las lecturas *EEG* obtenidas son accedidos por las clases definidas para el análisis y procesamiento, de forma que los datos de salida obtenidos conformen un conjunto de datos accesibles para el usuario final.

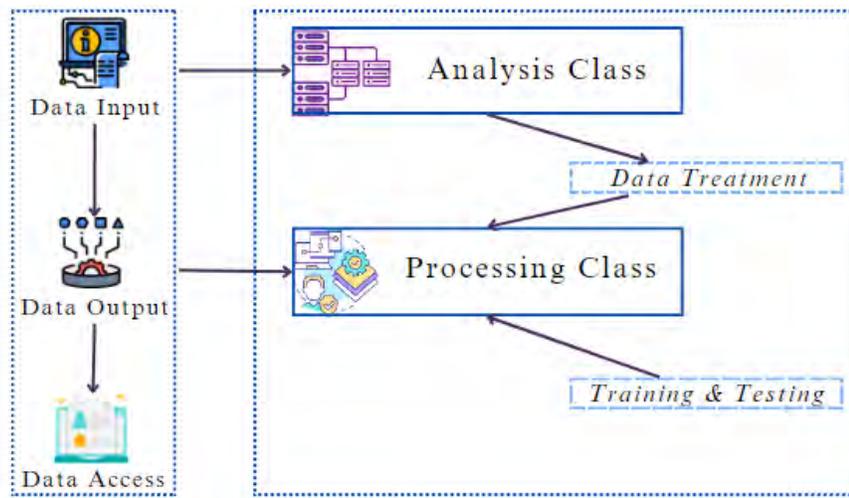


Figura 3.5: Diagrama de Contenedores del Sistema.

### 3.2.5. Evaluación: Simulación de Datos Locales

El prototipo del sistema fue replanteado con base a las interacciones del usuario en el sistema y los niveles de abstracción identificados en las secciones anteriores, consecuentemente estableciendo el primer escenario exploratorio de la metodología; la Figura 3.6 muestra la ejecución de *scripts* en Python para el análisis y procesamiento de componentes, en conjunto con el muestreo del conjunto de datos de bioseñales *EEG* a prueba y su desempeño.

```

    # FilterComponent.py
    import os
    import numpy as np
    import sine

    # Loading sample file that contains EEG and MEI data along with structural
    sample_data_folder = os.path.join(sample_data_path(),
    sample_data_raw_file = os.path.join(sample_data_folder, 'raw', 'sample_
    sample_data_raw_file')

    raw = sine.io.read_raw_fif(sample_data_raw_file)

    raw.crop(0, 0).load_data() # Use just 10 seconds of data, to save memory

    # Getting information regarding the structure of the loaded data
    print(raw)
    print(raw.info)

    # Introducing filtering by sine filter with given parameters
  
```

Figura 3.6: Evaluación del Sistema.

### 3.3. Diseño Experimental

En primera instancia, se realizó un análisis de la propuesta metodológica basada en *Lean UX* [56], así como sus cinco etapas principales: Segmentación para la exploración y observación de usuarios, Diseño que cubre prototipado y reuniones con expertos, Tratamiento para el conjunto de datos, implementación de técnicas de IA y procesos de *Big Data*, Lanzamiento que engloba pruebas de uso, y Evaluación que culmina el ciclo de pruebas con el usuario tal y como se ejemplificó en la Figura 3.1. Posteriormente, derivado de la etapa de tratamiento, se identificó el proceso para la implementación de las técnicas de IA partiendo de la Ciencia de Datos, mismo que se muestra en la Figura 3.7 conformado por las etapas de: Segmentación, Abstracción, Perfilado y Modelado de los datos; que se conforman de tres subetapas respectivamente: recolección, exploración y limpieza de los datos, análisis, tratamiento y reestructuración, evaluación, optimización y comparación, así como muestreo, entrenamiento e implementación.

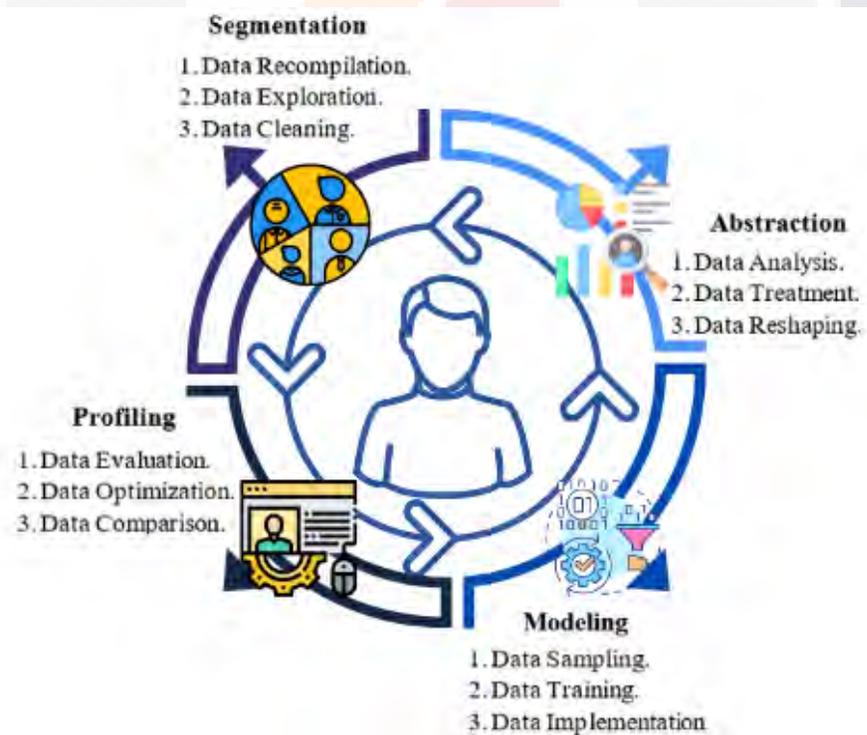


Figura 3.7: Contextualización de los Procesos de Ciencia de Datos e IA [30].

### 3.3.1. Segmentación de los Datos: Recolección, Exploración y Limpieza

La etapa de segmentación fungió como filtro para la identificación del conjunto de datos de referencia durante la etapa de experimentación local, y, por tanto, del tipo de datos a ser representados para posterior recolección y descarga. La contextualización de los procesos asociados de sus etapas consistió en:

- **Recolección.** Con el fin de obtener una segmentación adecuada de los datos por evaluar, se realizó una comparación entre conjuntos de datos y bases de datos de bioseñales *EEG* disponibles, considerando las magnitudes de *Big Data* [21]:
  - Velocidad. Capacidad de generación y procesamiento del conjunto de datos.
  - Variedad. Caracterización de datos heterogéneos.
  - Viabilidad. Consistencia en el comportamiento de los datos.
  - Veracidad. Integridad de la información representada.
  - Visualización. Interpretación de las características del conjunto de datos.
  - Valor. Factibilidad de hallazgos relevantes para la problemática.

Lo anterior, en concordancia con el enfoque a la solución de problemas de la Ciencia de Datos; se seleccionó un conjunto de datos para trabajar con algunas de las áreas de regiones cerebrales y sensores correspondientes [11, 45]. Véase Sección 4.2.1.

- **Exploración.** El tipo de datos de cada variable involucrada fue definido en conjunto con su clasificación y rango acorde a los valores y frecuencias registradas.
- **Limpieza.** Eliminación de variables innecesarias que contenían elementos fuera del rango, o valores nulos.

### 3.3.2. Abstracción de los Datos: Análisis, Tratamiento y Estructura

La abstracción involucró la preparación del conjunto de datos para que fueran procesados, consistente con los procesos de las subetapas:

- **Análisis.** Evaluación inicial de la estructura del conjunto de datos y clasificación de variables globales conforme a su denominación estadística y escala de medición [95]:
  - Variables Cuantitativas. Son la base de las métricas y se componen de datos numéricos; se clasifican en discretas y continuas.

- Discretas. Datos expresados como un todo, que se miden en números enteros.
- Continuas. Complemento de variables discretas que se expresa en decimales.
- Variables Cualitativas. Representación de las cualidades del conjunto de datos; se clasifican en nominales y ordinales.
  - Nominales. Denominación de nombres y propiedades de las variables.
  - Ordinales. Secuencia de valores representados de forma ordenada.

El conjunto de datos se definió como datos de tipo estructurado dado el contexto de la problemática; datos compuestos por variables cuantitativas asociadas a la lectura de bioseñales *EEG*. Véase Sección 4.2.2.

- **Tratamiento.** Transformación de las variables globales e identificación de parámetros para la implementación algorítmica del conjunto de datos; modificación de la nomenclatura de las columnas correspondiente a la ubicación del sensor en la región cerebral.
- **Reestructuración.** Normalización del conjunto de datos para la obtención de proyecciones, mediante una variable de control para clasificación y agrupación de estos.

### 3.3.3. Modelado de Datos: Muestreo, Entrenamiento e Implementación

Una vez completada la exploración del conjunto de datos, la transformación resultante fue tomada como referencia para la aplicación de técnicas de IA; las especificaciones y requerimientos fueron validados dados los datos por muestrear y entrenar, acorde a los procesos de las subetapas:

- **Muestreo.** Para que los datos fueran modelados congruentemente, se dio continuidad al Análisis Exploratorio de los Datos, o *EDA* por sus siglas en inglés; correspondiente a las pruebas y experimentos por realizar, con base al fundamento estadístico de las técnicas de IA seleccionadas y las funciones disponibles en el entorno de programación [80], cubriendo:
  - Muestra. Continuación del *EDA*, para profundizar el entendimiento sobre el conjunto de datos seleccionado.
  - Normalización. Regularización de los datos para mejorar el desempeño del modelo y ejecución de las técnicas de IA; aplicación de la función base de *scale()* disponible en el *software* de R Studio [132, 127].

- División. Observación de las tendencias del conjunto de datos respecto al *EDA* y normalización obtenida mediante *scale()* como criterios de *subsetting* para prueba, entrenamiento y validación de datos.

Con lo cual, derivado del proceso de *EDA* el conjunto de datos fue transformado en función del comportamiento de su modelo en las técnicas de IA. Véase Sección 4.2.3.

- **Entrenamiento.** Aplicación de la distribución parcial del conjunto de datos original, como subconjunto de Entrenamiento para los algoritmos seleccionados: *K-Means*, *ID3* y *J48*. Fundamento provisto por la validación del modelo de entrenamiento mediante los procesos de etapas anteriores como *ETL Pipelines* [128], siendo dependiente de las relaciones de los subconjuntos de la fuente de datos:

- Entrenamiento. Subconjunto central definido con base a la obtención de un radio de distribución representativa del conjunto de datos, de forma manual y mediante el uso de la función base de *split()* en el *software* de R Studio [132, 127].
- Validación. Subconjunto de control definido como métrica para el desempeño del modelo en el subconjunto de Entrenamiento.
- Prueba. Subconjunto de evaluación del comportamiento del subconjunto de Entrenamiento, tomando como control al subconjunto de Validación.

Subconjuntos definidos con base al *EDA* y tratamiento de los datos durante el Muestreo, conformando así los resultados obtenidos del Entrenamiento para los algoritmos en conjunto con su Validación y Prueba, sustentando el acondicionamiento del modelo con base a las directrices aportadas por *Machine Learning* [16] para los modelos de IA. Véase Sección 4.2.3.

- **Implementación.** Evaluación preliminar de desempeño del conjunto de datos original y el subconjunto de *Entrenamiento* para cada uno de los algoritmos, a través de su ejecución en el ambiente de *R Studio* [127], en preparación a su perfilado y validación.

### 3.3.4. Perfilado de Datos: Evaluación, Optimización y Comparación

Culminando con el diseño instrumental y el modelo basado en IA y Ciencia de Datos, se procedió con el perfilado del conjunto de datos [102]; tomando como parámetros la caracterización provista por el *EDA* y el Muestreo, en conjunto con las observaciones del comportamiento del modelo en los subconjuntos, abordando los procesos de las subetapas:

■ **Evaluación.** Los resultados obtenidos de la implementación del subconjunto de Entrenamiento fueron evaluados en concordancia a la caracterización provista para aprendizaje supervisado y no supervisado. Se adoptaron etapas propuestas del ciclo de vida de datos, para la usabilidad de los datos del mundo real [147], asociando los hallazgos obtenidos de la implementación de las técnicas de IA:

- **Adquisición.** Inspección del proceso de selección de la fuente de información, métodos de adquisición, tratamiento, modelado y transformación de la estructura del conjunto de datos; congruencia entre la estrategia propuesta para abordar la problemática partiendo de la segmentación de los datos, particularmente del proceso de Recolección, como nexos del Análisis efectuado para la abstracción de datos, y su modelado reflejado en el Muestreo del subconjunto de Entrenamiento.
- **Agregación.** Enriquecimiento del conjunto de datos mediante los procesos de Análisis y Muestreo, efectuando la reestructuración de los subconjuntos de datos mediante la redefinición de variables globales y asociación de los parámetros correspondientes al proceso de aprendizaje subsecuente.
- **Mantenimiento.** Exploración del flujo de datos para determinar los niveles de acceso como parte de la abstracción del modelo e identificación de las relaciones entre datos con base a su distribución inicial y los resultados obtenidos del *Entrenamiento*, así como la congruencia entre la distribución de la información en cada uno de los subconjuntos y la función a desempeñar.
- **Implementación.** Valoración de los hallazgos obtenidos como parte del ciclo de datos, puntualizando los valores de las métricas asociadas al desempeño del modelo y soluciones generadas por cada una de las técnicas de IA implementadas, así como la retroalimentación obtenida del nivel de acceso a datos y el flujo representado como comportamiento del modelo.

Por tanto, efectuando una revisión de la Entropía, Ganancia obtenida, y precisión de los árboles de decisiones *ID3* y *J48* [5], aunado al número de  $K$  obtenido con el Método de Codo para *K-Means* para la clasificación y agrupación de los datos como producto a iterar para su posible optimización. Véase Sección 4.2.4.

■ **Optimización.** Identificación de áreas de oportunidad en el aprendizaje supervisado, dado el comportamiento del modelo bajo el algoritmo *ID3* y sus limitantes en función del crecimiento exponencial de datos, siendo resultado de la inferencia, su sustitución por el algoritmo *J48*. Véase Sección 2.5.3.

- **Comparación.** Comparación del desempeño del algoritmo *ID3* y su versión optimizada *J48* en el modelo, validando la premisa de la Optimización en la precisión para la asignación de datos; aplicación de la técnica de *Model Fine Tuning* mediante la prueba y validación de los modelos generados en R Studio en paralelo a la implementación en Weka para *cross-validation*.



# TESIS TESIS TESIS TESIS TESIS

## Capítulo 4

### Resultados

#### 4.1. Contextualización Instrumental

Los resultados obtenidos de la investigación desarrollada tienen como base teórico-práctica la contextualización instrumental provista por el análisis de las técnicas algorítmicas mostradas en la Tabla 2.2; siendo producto directo de los Componentes para Análisis y Procesamiento mostrados en la Figura 3.4, la implementación de los algoritmos de *K-Means*, *ID3* y *J48* cuyo *output* resultante de los procesos orientados al Tratamiento de Datos, así como Entrenamiento y Pruebas integradas en el esquema de Contenedores ejemplificado en la Figura 3.5. Cabe denotar, que acorde al preámbulo de la Metodología introducida en la Sección 3.1, y el Modelo de Ciencia de Datos descrito en la Sección 3.3 como *MVP* clave de la iteración metodológica, la implementación de las técnicas de IA culminó en el mapeo de los *MVPs* diseñados, por medio de su integración al Modelo de Ciencia de Datos de la Figura 3.7, bajo el esquema de 12 procesos derivados de las subetapas que le componen.

##### 4.1.1. Implementación del Algoritmo K-Means

Siguiendo los pasos delimitados para el Pseudocódigo del algoritmo *K-Means* mostrado en la Figura 2.2, la implementación del algoritmo consistió en:

- **Determinar K.** Uso del método gráfico para observar el comportamiento del conjunto de datos, aplicando el Método de Codo introducido en la Sección 2.5.3, definiendo el número de grupos factibles e identificando el Hiperparámetro para su estudio dado por la Ecuación 2.3.

- **Posicionar Centroides.** Carga del conjunto de datos al ambiente de R Studio, inicializando los centroides en coordenadas aleatorias del espacio de solución.
- **Asignar Datos a K.** Cálculo de la distancia media entre los puntos del conjunto de datos y sus centroides, aplicando la Ecuación 2.4 para el uso de la Distancia Euclidiana.
- **Iterar la Asignación.** Repetición del proceso para asignación de puntos al clúster más cercano, hasta que todos los elementos se encontraron asignados y no hubo cambio de clúster u posicionamiento de los centroides, en adhesión a la función objetivo de minimizar y la Ecuación 2.2.

En lo que respecta a la identificación del número de grupos factibles  $K$  mediante la aplicación del Método de Codo, los resultados de su implementación en el *software* de R Studio fueron corroborados con los parámetros de la herramienta de Weka para *K-Means*, obteniéndose un valor congruente de  $K=2$  con el mejor desempeño, lo cual puede observarse en la Figura 4.1 dada la característica curvatura del método, donde el número de clústeres es representado en el eje X, en función de los valores de WSS en el eje Y:

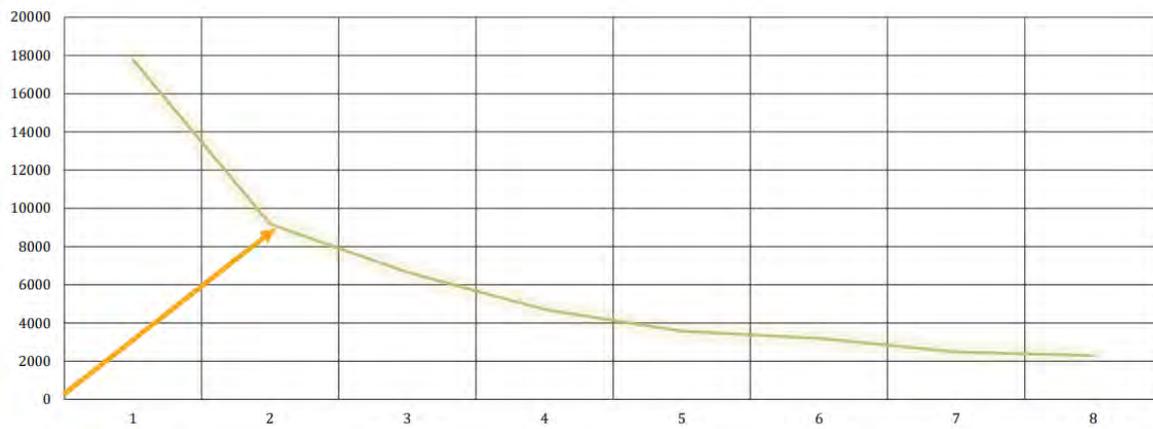


Figura 4.1: Determinación de K con el Método de Codo.

En concordancia con el *EDA* descrito en la Sección 3.3.3, se procedió a generar un modelo de regresión lineal de los datos en crudo, cuyo gráfico de dispersión permitió evaluar el comportamiento del algoritmo de *K-Means* y su precisión para el agrupamiento en clústeres, corroboración efectuada con el cálculo de su número idóneo  $K$ , dado por la Ecuación 2.3 de la *WSS*; los puntos de datos fueron distribuidos en dos clústeres con **67 %** y **17.9 %** de los datos respectivamente, como se muestra en la Figura 4.2:

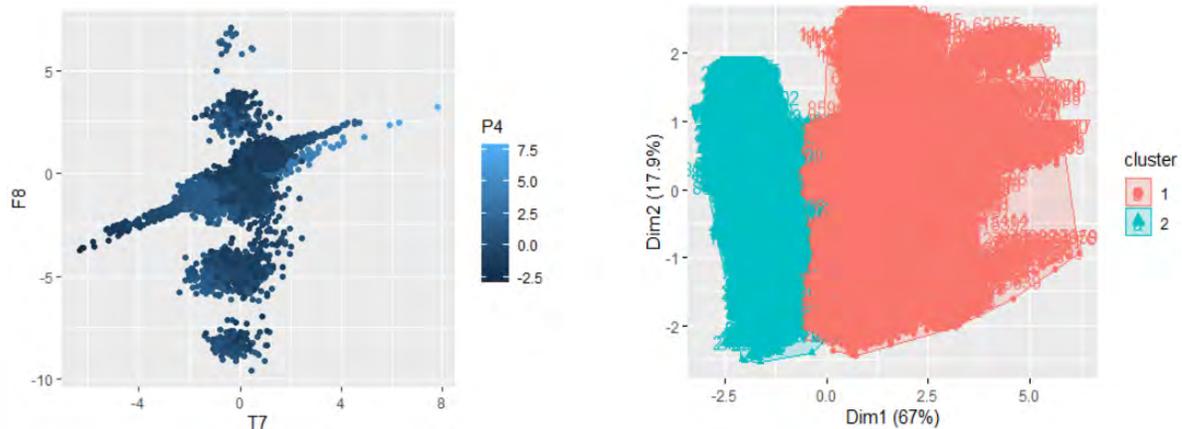


Figura 4.2: Dispersión del Conjunto de Datos y Clusterización.

Posteriormente, el porcentaje fue ajustado a **57 %** y **43 %** tras corroborarse la distribución de los puntos de datos en la herramienta de Weka y transformar el conjunto de datos, iterando en el Modelo de Ciencia de Datos los procesos derivados de la Abstracción y Modelado; habiéndose obtenido como hallazgo, la importancia de los registros **E4** y **E6** para la tendencia de agrupación del conjunto de datos.

#### 4.1.2. Implementación de los Algoritmos ID3 y J48

Como complemento a los resultados obtenidos mediante el aprendizaje no supervisado de *K-Means*, se procedió a categorizar el conjunto de datos mediante la serie de pasos provistos por el Pseudocódigo del algoritmo *ID3* mostrado en la Figura 2.3, cubriendo:

- **Determinar el Nodo Raíz.** Análisis del conjunto de datos para validar su conformación por subconjuntos positivos y negativos en función de la Ecuación 2.5; seleccionando el atributo con mayor peso como nodo raíz *A*, aplicando la Ecuación 2.6.
- **Evaluar Atributos.** Segmentación de las relaciones del conjunto de datos representadas en proporción a la Ganancia esperada de Entropía acorde a la Ecuación 2.7.
- **Asignar Categorías.** Categorización del conjunto de datos en la estructura del árbol como hojas o subárboles de la estructura original.
- **Iterar la Asignación.** Ejecución del proceso para asignación de nodos, hojas y ramas hasta obtener la mayor Ganancia para *A*.

El panorama de la agrupación de datos fue esencial para la aplicación del *EDA* como acercamiento a la identificación de métricas, aplicación de métodos de exploración, observación, análisis y tratamiento como proyección del comportamiento esperado del modelo; puntualmente, la implementación del algoritmo *ID3* permitió desglosar los atributos con mayor peso dentro del conjunto de datos así como las relaciones existentes entre estos en medida de la Ganancia y Entropía que poseen, como puede observarse en el árbol de clasificación obtenido que se muestra en la Figura 4.3 con las asignaciones correspondientes por categoría acorde a lo identificado a través de la implementación del algoritmo de *K-Means*, en que si bien los registros **E4** y **E6** no resultaron elegidos como un nodo raíz de primer o segundo orden, son en conjunto con los registros de **E9** los atributos con mayor ganancia de información.

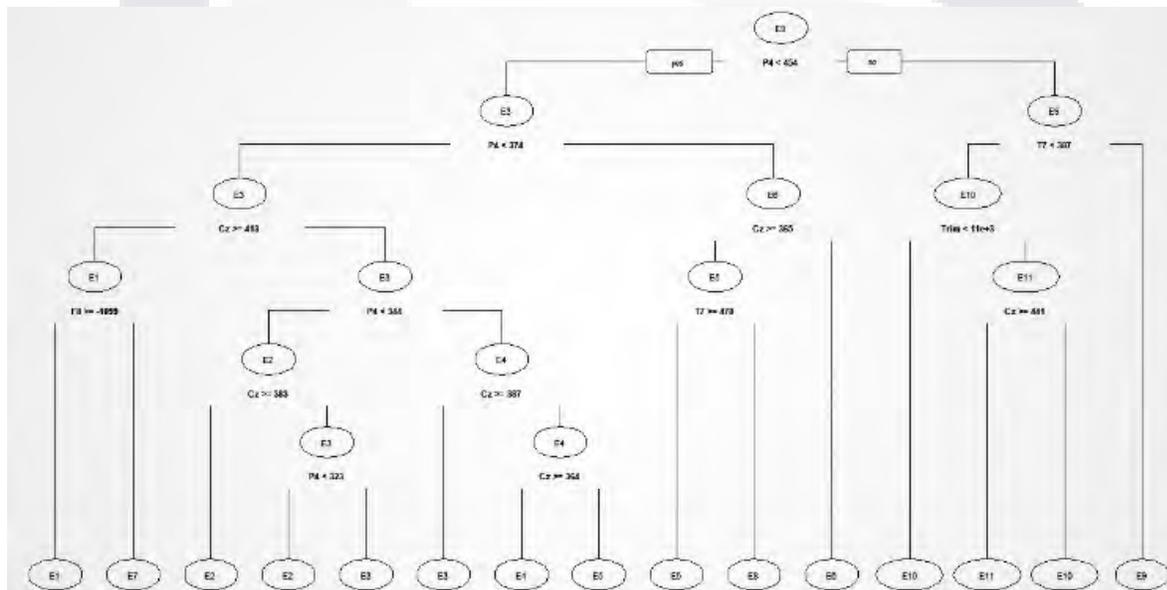


Figura 4.3: Árbol de Clasificación *ID3*.

Si bien el volumen de datos corresponde a un total de 144, 000 registros, el árbol solo tuvo que ser recortado en un nodo de tercer nivel, y las 12 categorías producidas coincidieron con el total de variables globales de la fuente de datos de origen, por lo que la agrupación precedente no incurrió en *overfitting* del modelo. Adicionalmente, en cuanto los pesos de los datos categóricos fueron obtenidos, para corroborar los hallazgos iniciales se procedió con la ejecución de otra prueba para evaluar el desempeño del modelo de entrenamiento, ello mediante la generación de Matrices de Confusión de los algoritmos *ID3* y *J48*, que permitió evaluar la precisión del modelo y tendencias observadas para *ID3*, obteniéndose de **50 %** a **80 %** de precisión, mismo que tras evaluar la adquisición basada en Entropía, sustentó su

reemplazo en el modelo final por su versión optimizada: el algoritmo *J48*; por consiguiente, los resultados finales de la implementación del modelo arrojaron una precisión de **89 %** a **96 %**, lo cual puede observarse en la Figura 4.4:

	E1	E10	E11	E2	E3	E4	E5	E6	E7	E8	E9	
E1	9706	0	0	956	29	0	0	0	0	0	0	
E10	0	11603	2097	20	6	0	14	142	0	130	1279	
E11	0	2335	5719	4	9	0	9	27	0	102	486	
E2	1600	0	0	12074	1026	31	130	0	0	0	30	
E3	64	0	0	1705	12895	2376	238	5	3	4	1	
E4	0	0	0	0	1190	12056	3443	1	0	0	1	
E5	0	0	0	1	38	1060	10877	1713	0	2	0	
E6	0	0	0	0	2	0	1558	14129	0	1	1	
E7	3	0	0	7	0	0	33	0	6410	38	0	
E8	0	154	35	0	8	1	25	0	183	4428	1657	
E9	0	723	26	0	0	0	0	0	0	63	17278	

	a	b	c	d	e	f	g	h	i	j	k	classified as
17908	0	0	0	68	0	0	0	17	0	97	a = E9	
0	16428	354	3	0	466	32	2	0	1	5	b = E3	
0	268	16050	2	0	4	366	0	0	0	1	c = E4	
0	3	3	15596	1	0	87	0	0	0	1	d = E6	
74	0	0	0	14325	0	0	0	874	0	18	e = E10	
0	531	4	0	0	14213	0	143	0	0	0	f = E2	
0	15	373	70	0	0	13233	0	0	0	0	g = E5	
0	0	0	0	0	165	0	10526	0	0	0	h = E1	
40	0	0	0	897	0	0	0	7741	0	13	i = E11	
0	4	0	0	0	0	0	4	0	6477	6	j = E7	
136	0	1	2	44	1	2	0	15	4	6286	k = E8	

Figura 4.4: Matrices de Confusión de *ID3* y *J48*.

Que respalda la importancia de los ajustes realizados, con el algoritmo *ID3*, cuya Entropía coadyuvó a la segmentación de relaciones, haciéndolo un punto de comparación para identificar el comportamiento del conjunto de datos y sus tendencias, mientras que su versión optimizada, el algoritmo *J48* aportó la optimización requerida para la adquisición de hallazgos a través de fuentes de información existentes y ejecución de procesos para el modelado de datos en función del incremento en Entropía como factor diferencial de la Validación y Verificación de los resultados obtenidos del proceso iterativo.

## 4.2. Asignación de Procesos de IA

Los resultados obtenidos a través de la primera iteración del diseño metodológico basado en *Lean UX* en conjunto con el enfoque de Ciencia de Datos, cubrieron las bases para los procesos de Validación y Verificación del desarrollo de *software*, al igual que el *EDA* del modelo para futuras aplicaciones respecto a la correlación entre bioseñales *EEG* y el mapeo de regiones cerebrales para la salud mental.

La Figura 4.5 muestra las ramificaciones de los resultados obtenidos, en que el enfoque de *Lean UX* [56] es representado como el núcleo de la metodología, tomando en consideración los *MVPs* que son obtenidos tras cada una de las etapas de la metodología, y priorizando los procesos de Verificación y Validación en su última etapa, mientras que el enfoque de Ciencia de Datos es representado como el proceso de diseño que engloba los avances actuales presentados por medio de la metodología de Ciencia de Datos propuesta por *IBM* [118], y su foco en la problemática, siendo que esta fue particularmente explorada en la comparación y selección de técnicas algorítmicas para la categorización y agrupación de datos.

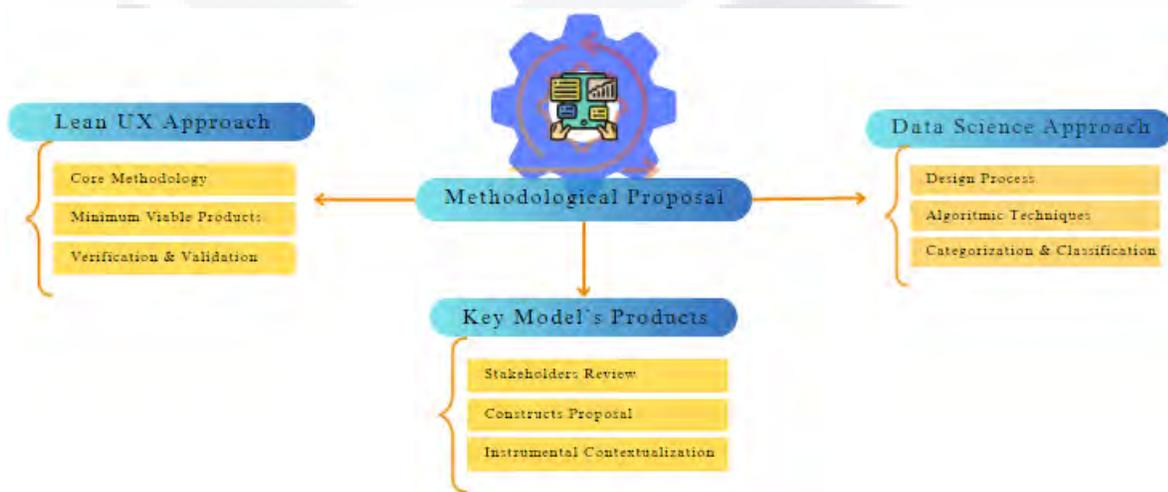


Figura 4.5: Análisis Exploratorio del Modelo.

Descritas en las subsecciones que se muestran a continuación, se encuentran las etapas identificadas como parte clave del Modelo de Ciencia de Datos y por consiguiente de la etapa de Tratamiento englobada en la propuesta metodológica, en conjunto con sus contribuciones al proceso de diseño para la correlación de bioseñales *EEG* a la par de la identificación de los productos principales del modelo: la revisión con *stakeholders*, la propuesta de los constructos, y la contextualización instrumental. A su vez, como parte intrínseca del enfoque iterativo de la metodología, en el Anexo A puede observarse el desglose puntual de cada uno de los productos obtenidos por iteración.

#### 4.2.1. Segmentación: Variables de Estudio Globales

Tal y como se estableció en el Diseño Experimental de la Metodología, la Segmentación de los datos fue utilizada como filtro para el conjunto de datos a explorar mediante la identificación de variables involucradas y sus efectos. Véase Sección 3.3.1.

Partiendo del análisis de la estructura de datos en crudo como un conjunto de elementos a explorar sin sesgar la información para el diseño del instrumento orientado a la ansiedad; el conjunto de datos de *Auditory Evoked Potential* [11] fue elegido como fuente de datos a segmentar, adicional a su compatibilidad con el dispositivo de medición de señales *EEG* para comparar futuras referencias, y la colocación de los sensores en las regiones cerebrales: frontal *F8*, parietal *P4*, y temporal *Tz* [45]. Como se muestra en la estructura de datos propuesta para la delimitación de columnas en la Figura 4.6:

	Trim	Control	T7	F8	P4	Cz
56486	11409	E3	516.9832	-986.5071	351.4926	399.9337
56487	11409	E4	503.5214	-1000.1278	353.6823	370.9139
56488	11409	E5	502.5079	-980.7289	373.6048	372.5875
56489	11409	E6	494.8000	-984.0444	398.0899	355.5822
56490	11409	E9	401.8242	-1237.4451	474.7579	431.0721
56491	11410	E1	505.7130	-1015.8616	320.8441	413.4347
56492	11410	E10	378.5919	-1260.8960	501.0811	437.5253
56493	11410	E11	366.7458	-1262.4668	507.3230	442.0468
56494	11410	E2	486.6451	-1009.1559	313.0577	386.8440
56495	11410	E3	495.9276	-1007.3196	332.5949	378.3956
56496	11410	E4	496.6999	-1010.5023	355.3429	371.8340
56497	11410	E5	520.6577	-976.0578	379.5999	376.4284
56498	11410	E6	496.4904	-986.2434	400.2441	354.3406
56499	11410	E9	395.5748	-1247.0079	463.4147	419.0595

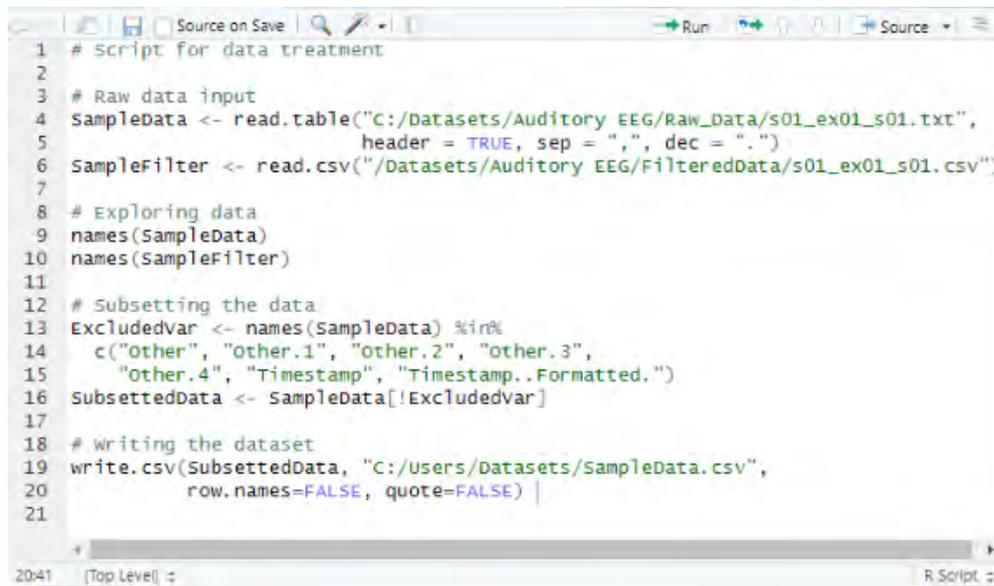
Showing 56,486 to 56,499 of 286,000 entries, 6 total columns

Figura 4.6: Variables de Estudio del Conjunto de Datos.

### 4.2.2. Abstracción: Transformación del Espacio Experimental

Con el conjunto de datos establecido y sus parámetros definidos en el análisis preliminar, la implementación de la etapa de Abstracción permitió preparar el conjunto de datos para su posterior evaluación. Véase Sección 3.3.2.

La reestructuración propuesta en la Figura 4.6 se aplicó a una muestra aleatoria del conjunto de datos, tras un análisis simultáneo con datos previamente filtrados, posterior a lo cual se efectuó el Tratamiento y Reestructuración de los datos vía *subsetting*, para que estos fueran descriptivos; ejecutando el *script* en *R* que se muestra en la Figura 4.7:



```

1 # Script for data treatment
2
3 # Raw data input
4 SampleData <- read.table("C:/Datasets/Auditory EEG/Raw_Data/s01_ex01_s01.txt",
5 header = TRUE, sep = ",", dec = ".")
6 SampleFilter <- read.csv("/Datasets/Auditory EEG/FilteredData/s01_ex01_s01.csv")
7
8 # Exploring data
9 names(SampleData)
10 names(SampleFilter)
11
12 # Subsetting the data
13 ExcludedVar <- names(SampleData) %in%
14 c("other", "other.1", "other.2", "other.3",
15 "other.4", "Timestamp", "Timestamp..Formatted.")
16 SubsetData <- SampleData[!ExcludedVar]
17
18 # writing the dataset
19 write.csv(SubsetData, "C:/Users/Datasets/SampleData.csv",
20 row.names=FALSE, quote=FALSE) |
21

```

Figura 4.7: *Script* de Tratamiento de Datos en R.

Habiendo efectuado el tratamiento por ser aplicado en conjunto con la normalización de variables y parámetros, que validan las proyecciones y cálculos de las técnicas algorítmicas, por ser implementadas para la clasificación y agrupación de variables de control en los subconjuntos de datos de Entrenamiento, Validación y Prueba.

#### 4.2.3. Modelado de Datos: Implementación del Modelo de IA

La transformación final del conjunto de datos fue cotejada para cada una de las técnicas algorítmicas y las variables de control a ser consideradas para los cambios presentes entre el conjunto de datos en crudo y los datos transformados, resultando en un conjunto de datos de muestra, mismo que fue entrenado durante la etapa de experimentación acorde a las especificaciones provistas para cada tipo de algoritmo y categorías resultantes. Véase Sección 3.3.3.

La implementación del *script* que se muestra en la Figura 4.8 realizó el producto de los resultados parciales y los hallazgos provistos por la aplicación de las técnicas algorítmicas, haciendo factible el perfilado de los datos para implementaciones futuras, así como el acondicionamiento del modelo acorde a las 12 categorías obtenidas a través de *K-Means*, *ID3* y *J48* correspondientes a la experimentación para el conjunto de datos de la población muestra.



```

3 # Each file in the raw data contains additional columns that can be discarded.
4 D1 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex01_s01.csv")
5 D2 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex01_s02.csv")
6 D3 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex01_s03.csv")
7 D4 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex02_s01.csv")
8 D5 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex02_s02.csv")
9 D6 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex02_s03.csv")
10 D7 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex05.csv")
11 D8 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex06.csv")
12 D9 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex07.csv")
13 D10 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex08.csv")
14 D11 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex09.csv")
15 D12 <- read.csv("Datasets/Auditory EEG/Segmented_Data/s01_ex10.csv")
16
17 # Subsetting the data to change colnames, add one for control, and shift them
18 names(D1); names(D1)[1] = "Trim"; set.seed(6); Control <- sample(c("E1"), size
19 names(D2); names(D2)[1] = "Trim"; set.seed(6); Control <- sample(c("E2"), size
20 names(D3); names(D3)[1] = "Trim"; set.seed(6); Control <- sample(c("E3"), size
21 names(D4); names(D4)[1] = "Trim"; set.seed(6); Control <- sample(c("E4"), size
22 names(D5); names(D5)[1] = "Trim"; set.seed(6); Control <- sample(c("E5"), size
23 names(D6); names(D6)[1] = "Trim"; set.seed(6); Control <- sample(c("E6"), size
24 names(D7); names(D7)[1] = "Trim"; set.seed(6); Control <- sample(c("E7"), size
25 names(D8); names(D8)[1] = "Trim"; set.seed(6); Control <- sample(c("E8"), size
26 names(D9); names(D9)[1] = "Trim"; set.seed(6); Control <- sample(c("E9"), size
27 names(D10); names(D10)[1] = "Trim"; set.seed(6); Control <- sample(c("E10"), siz
28 names(D11); names(D11)[1] = "Trim"; set.seed(6); Control <- sample(c("E11"), siz
29 names(D12); names(D12)[1] = "Trim"; set.seed(6); Control <- sample(c("E12"), siz
30
31 # Merging the datasets
32 DFList <- list(D1, D2, D3, D4, D5, D6, D7, D8, D9, D10, D11, D12); SubsetData
33
34 # Writing the dataset
35 write.csv(SubsetData, "Datasets/MergedData.csv", row.names=FALSE, quote=FALSE

```

Figura 4.8: *Script* de Reestructuración de Datos en R.

#### 4.2.4. Perfilado: Evaluación e Iteración del Modelo

El proceso de diseño para la metodología y aspectos clave del modelo coadyuvaron a establecer el diseño instrumental de las técnicas computacionales y su aplicación para el perfilado de los datos provistos, por tal motivo la contribución a la evaluación y optimización de los procesos de la metodología y el modelo mismo hicieron del análisis de los resultados parciales y los *MVPs* un factor de decisión para comparar el nivel de progreso respecto a la implementación del proyecto y lo idóneo del conjunto de datos así como soluciones provistas orientadas a la correlación de ansiedad y sus indicadores. Véase Sección 3.3.4

Como se muestra abajo, la Figura 4.9 introduce las etapas del proceso de diseño acorde a su función principal, subrutinas por desarrollar, y constructos identificados para su diseño dado el involucramiento de los *stakeholders* en el proceso, complementando la metodología propuesta en la Figura 3.1 y el análisis exploratorio en la Figura 4.5 al proveer puntos de inflexión que reforzaran la investigación reproducible y se convertirán en el enlace entre las bases de la Ingeniería de Software, y el diseño y aplicación de la instrumentación con las técnicas identificadas.

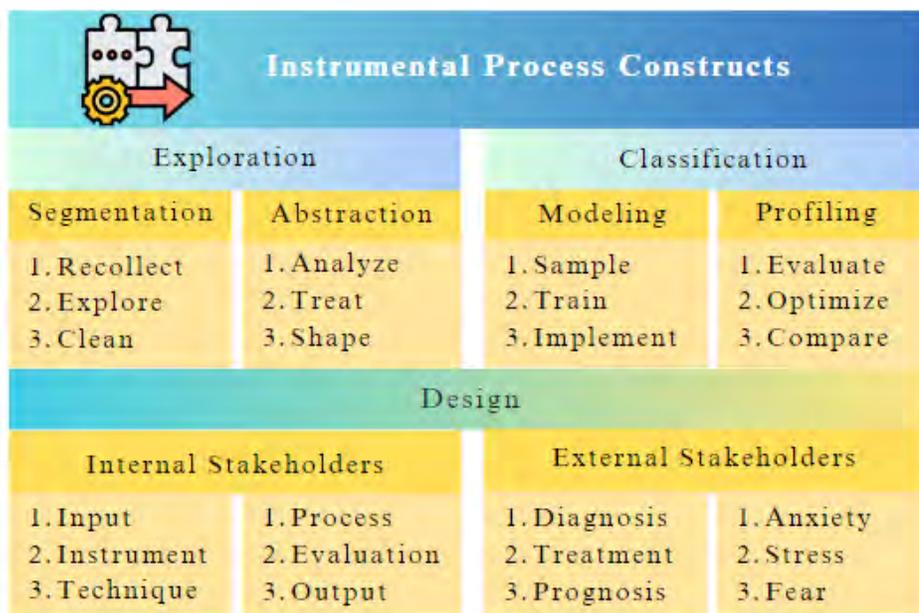


Figura 4.9: Diseño del Proceso Instrumental.

Aunado al desglose de los resultados obtenidos, el Anexo B muestra el impacto de las publicaciones realizadas como parte de los productos obtenidos para la generación de resultados de la metodología.

# TESIS TESIS TESIS TESIS TESIS

## Capítulo 5

### Conclusiones

El diseño holístico y análisis de la metodología propuesta en la Sección 3.1 así como las etapas y procesos que la conforman, permitieron contrastar la hipótesis formulada con el propósito de su diseño y objetivos de investigación, cuyas generalidades delimitadas en el Capítulo 1, en conjunto con el análisis de los requerimientos del usuario mediante la Guía de Diseño de *MVPs* planteada en la Sección 3.2 proveyeron directrices para la reproducción del modelo generado e implementación del *feedback* como parte del ciclo de mejora continua de la metodología; subsecuentemente, los procesos en torno a los *MVPs* delimitaron la importancia de su capacidad de producción así como escalabilidad al considerarse el crecimiento exponencial para el esquema de desarrollo de *software* y aplicación del enfoque de *Lean UX*, denotando el requerimiento de un conjunto de estándares de diseño para las etapas de prototipado y *feedback*, con el objeto de documentar el grado de avance en lo referente al prototipado, así como una métrica asociada al *feedback* provisto para cuantificar la satisfacción del usuario e integrar los resultados a las métricas correspondientes de la intervención con el especialista de la salud mental, particularmente de ansiedad y ataques de pánico.

Aunado a lo anterior, el Diseño Experimental presentado en la Sección 3.3 y los datos obtenidos de los procesos designados que pasaron por las pruebas, concedieron el medio para la aplicación de los algoritmos al análisis y procesamiento de datos, cubriendo así el contexto de aplicación y asignación mediante las abstracciones del sistema planteadas en la Figura 3.4 y la Figura 3.5; ello, al estar orientados a la obtención de consultas asociadas al nivel de equipamiento requerido para el monitoreo de la actividad, haciendo factible la calibración del dispositivo *EEG*, el procesamiento y análisis de la señal sin delimitantes considerables consecuentes del ruido en las señales; siendo pertinente la aplicación de algoritmos basados en el conjunto de datos como parte de los procesos centrales de la metodología y su modelo, para lo cual fue fundamental el análisis realizado con base a las técnicas de aprendizaje supervi-

sado y no supervisado que se ilustró en la Tabla 2.2, con la inteligencia artificial fungiendo como nexo entre las áreas de aplicación de la neurociencia y las directivas provistas por la metodología, potenciando los modelos subyacentes. Por otra parte, la problemática y caso de estudio deberán permanecer como la prioridad principal para evaluar el grado de avance hacia el objetivo de la investigación y proyecciones de los resultados, no solo como contraste hacia la hipótesis inicial y medida de las capacidades de estudio para trabajos relacionados o replicación de estudios específicos, sino para la implementación del enfoque analítico de forma iterativa tras cada avance parcial, ejecución u evaluación para aplicar el *feedback* obtenido de la última iteración al momento, o bien del registro de los *MVPs* y las proyecciones establecidas; por ello, incluso cuando la metodología de *Lean UX* es la que provee el *framework* para el modelo basado en Ciencia de Datos y los procesos basados en datos, los principios de estos se encuentran integrados como parte de la filosofía de *Lean UX*, connotado a través de la transparencia de los datos en los resultados obtenidos así como capacidad exponencial para la mejora, optimización y escalabilidad, conllevando a la correlación entre señales *EEG* y eventual mapeo de las regiones cerebrales en que se presenta ansiedad, con la intervención de especialistas de la salud mental.

Puede entonces concluirse que los lineamientos obtenidos de los procesos diseñados bajo la metodología propuesta condujeron a la asignación de tareas correspondientes a sus etapas dentro del flujo de datos, resaltando las implicaciones de las técnicas algorítmicas implementadas como parte de la Contextualización Instrumental y Asignación de Procesos de IA desglosados en el Capítulo 4, así como el procesamiento de los resultados hacia un perfil unificado que pueda ser modelado mediante el análisis y procesamientos de los datos asociados a este para la obtención de información complementaria al dispositivo *EEG*, los componentes clave de la *BCI*, las variables categóricas como control en la población muestra, y las mejoras por ser implementadas; siendo parte de los trabajos a futuro los hallazgos obtenidos de la experimentación local y la reevaluación del diseño metodológico que culminó en la introducción del modelo de ciencia de datos, mismos que deberán ser ajustados para sostener reuniones con especialistas de la salud mental para la identificación de la instrumentación por emplear en conjunto con el dispositivo *EEG* para la obtención de datos de los usuarios objetivo. Cabe mencionarse que la metodología propuesta para el diseño experimental de la correlación entre señales *EEG* y regiones cerebrales fue complementada con el proceso de diseño basado en Ciencia de Datos y técnicas de IA dentro de las restricciones de tiempo y recursos disponibles para el desarrollo del proyecto de investigación, con los resultados obtenidos como parte del diseño experimental focalizando el estudio del comportamiento del conjunto de datos en crudo, brindando puntos destacados respecto a los ajustes requeridos

no solo para la limpieza y análisis de los datos, sino para su procesamiento e interpretación conformando los resultados en cada uno de los algoritmos aplicados como parte del Perfilado de Datos desglosado en la Sección 3.3.4, con *K Means* y el número óptimo de clústeres obtenidos como la base sobre la cual el árbol de decisiones y la asignación de porcentajes del algoritmo *J48* fue implementado como mejora del desempeño provisto por el algoritmo *ID3*; sucesivo a la correspondiente clasificación y agrupación que podrán ser empleadas para las etapas de lanzamiento y evaluación de la metodología, particularmente en el perfilado y modelado de los datos, aportando al diseño y verificación del instrumento para la obtención de datos de pacientes que han tenido o presentan ansiedad, y como tal estableciendo muestras de entrenamiento y control de la población establecida, focalizando ansiedad, ansiedad generalizada, y la presencia y factibilidad de presentar ataques de pánico.

De lo anterior, que la continuidad del proyecto de investigación propuesto cubra las etapas tempranas de implementación del instrumento y los resultados de un conjunto de entrenamiento para ser comparado en etapas subsecuentes de la recolección de los datos una vez que el instrumento se encuentre terminado, de manera tal que el conjunto de datos de control y entrenamiento así como el flujo de datos en crudo pueda ser analizado, comparado y procesado para explorar la correlación entre las bioseñales *EEG* y las regiones cerebrales, mientras que se proporciona un lineamiento para los constructos objetivo durante las etapas de pruebas con el usuario. Véase el Anexo C.

## Referencias

- [1] David Aarlién y Ricardo Colomo-Palacios. «Lean ux: A systematic literature review». En: *International Conference on Computational Science and Its Applications*. Springer. 2020, págs. 500-510.
- [2] Reem Abdalla y Alok Mishra. «Agent-Oriented Software Engineering Methodologies: Analysis and Future Directions». En: *Complexity* 2021 (dic. de 2021), págs. 1-21. DOI: 10.1155/2021/1629419.
- [3] United States. National Aeronautics and Space Administration. Technology Utilization. *Biomedicine; a compilation*. [Washington], 1973. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edshtl&AN=edshtl.MIU01.008163426&lang=es&site=eds-live&scope=site>.
- [4] Zeeshan Ahmed et al. «Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine». En: *Database: The Journal of Biological Databases and Curation* 2020 (2020). ISSN: 17580463. DOI: 10.1093/DATABASE/BAAA010.
- [5] Faisal Mashel Albagmi et al. «Prediction of generalized anxiety levels during the Covid-19 pandemic: A machine learning-based modeling approach». En: *Informatics in Medicine Unlocked* 28 (ene. de 2022), pág. 100854. ISSN: 2352-9148. DOI: 10.1016/J.IMU.2022.100854.
- [6] Patricia Valle de Albuquerque Lima. «Gestalt-terapia e Cuidado.» En: *Estudos e Pesquisas em Psicologia* 19.4 (2019), págs. 1051-1066. ISSN: 16763041. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=asn&AN=143606309&lang=es&site=eds-live&scope=site>.

- [7] Ricardo Aler, Inés M. Galván y José M. Valls. «Transition Detection for Brain Computer Interface Classification». En: *Biomedical Engineering Systems and Technologies*. Ed. por Ana Fred, Joaquim Filipe y Hugo Gamboa. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, págs. 200-210. ISBN: 978-3-642-11721-3.
- [8] Bustamante Alexánder Rodríguez et al. «Terapia familiar sistémica. Sobre el self del terapeuta.» En: *Revista Latinoamericana de Estudios de Familia* 11 (2019), págs. 89-108. ISSN: 2145-6445. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsdoj&AN=edsdoj.96028a74d554f1fa4f4b50fc6cf2a5e&lang=es&site=eds-live&scope=site>.
- [9] Mohamed Alloghani et al. «A systematic review on supervised and unsupervised machine learning algorithms for data science». En: *Supervised and unsupervised learning for data science* (2020), págs. 3-21.
- [10] Fahad H Alshammari. «Trends in Intelligent and AI-Based Software Engineering Processes: A Deep Learning-Based Software Process Model Recommendation Method.» En: *Computational intelligence and neuroscience 2022* (2022), pág. 1960684. ISSN: 1687-5273. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=36248934&lang=es&site=eds-live&scope=site>.
- [11] Nibras Abo Alzahab et al. «Auditory evoked potential EEG-Biometric dataset». En: (2021). DOI: 10.13026/ps31-fc50. URL: <https://physionet.org/content/auditory%5C-eeg/1.0.0/>.
- [12] Nait-Ali Amine. *Advanced Biosignal Processing*. Springer, 2009. ISBN: 9783540895053. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=277209&lang=es&site=eds-live&scope=site>.
- [13] IBM Analytics. *Metodología Fundamental para la Ciencia de Datos*. [Online]. 2015. URL: <https://www.ibm.com/downloads/cas/6RZMKDN8>.
- [14] Howard Anton y Chris Rorres. *Elementary linear algebra: applications version*. John Wiley & Sons, 2013.
- [15] Violeta Argudo Portal. «Biomedicina en proceso de elaboración: Un estudio cualitativo sobre los biobancos como infraestructuras para la investigación biomédica.» Tesis de maestría. Universitat Autònoma de Barcelona, 2021., 2021. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?>

direct=true&db=edstdx&AN=edstdx.10803.674032&lang=es&site=eds-live&scope=site.

- [16] *ARSET - Fundamentals of Machine Learning for Earth Science | NASA Applied Sciences*. URL: <https://appliedsciences.nasa.gov/join-mission/training/english/arset-fundamentals-machine-learning-earth-science>.
- [17] Afshin Ashofteh y Jorge M. Bravo. «Data science training for official statistics: A new scientific paradigm of information and knowledge development in national statistical systems». En: *Statistical Journal of the IAOS* 37 (3 2021), págs. 771-789. ISSN: 18747655. DOI: 10.3233/SJI-210841.
- [18] American Psychiatric Association et al. «DSM-5: Manual diagnóstico y estadístico de los trastornos mentales». En: (2014).
- [19] Annushree Bablani et al. «A multi stage EEG data classification using k-means and feed forward neural network». En: *Clinical Epidemiology and Global Health* 8 (3 sep. de 2020), págs. 718-724. ISSN: 2213-3984. DOI: 10.1016/J.CEGH.2020.01.008.
- [20] Oana Bălan et al. «An investigation of various machine and deep learning techniques applied in automatic fear level detection and acrophobia virtual therapy». En: *Sensors* 20.2 (2020), pág. 496.
- [21] Kornelia Batko y Andrzej Ślęzak. «The use of Big Data Analytics in healthcare». En: *Journal of big Data* 9.1 (2022), pág. 3.
- [22] Morgan Beurenaut et al. «The ‘Threat of Scream’ paradigm: a tool for studying sustained physiological and subjective anxiety». En: *Scientific reports* 10.1 (2020), págs. 1-11.
- [23] Silas Bergen et al. «Classifying behavior from short-interval biologging data: An example with GPS tracking of birds». En: *Ecology and Evolution* 12 (2 feb. de 2022), e08395. ISSN: 2045-7758. DOI: 10.1002/ECE3.8395. URL: <https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.8395>.
- [24] Michael W Berry, Azlinah Mohamed y Bee Wah Yap. *Supervised and unsupervised learning for data science*. Springer, 2019.

- [25] J.J. Bird et al. «Synthetic Biological Signals Machine-Generated by GPT-2 Improve the Classification of EEG and EMG Through Data Augmentation.» En: *IEEE Robotics and Automation Letters, Robotics and Automation Letters, IEEE, IEEE Robot. Autom. Lett* 6.2 (2021), págs. 3498-3504. ISSN: 2377-3766. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsee&AN=edsee.9345373&lang=es&site=eds-live&scope=site>.
- [26] Carolina Blaya et al. «Lack of association between the Serotonin Transporter Promoter Polymorphism (5-HTTLPR) and Panic Disorder: a systematic review and meta-analysis». En: *Behavioral and brain functions* 3.1 (2007), págs. 1-10.
- [27] Hans-Hermann Bock. *Clustering Methods: A History of k-Means Algorithms*. Springer Berlin Heidelberg, 2007. ISBN: 978-3-540-73558-8. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.30502E1B&lang=es&site=eds-live&scope=site>.
- [28] Mikolaj Buchwald, Aleksandra Czyz y Marcin Jukiewicz. «Analysis of EEG Signals for Non-Technical and Non-Medical Students.» En: *International Journal of Instruction* 14.2 (2021), págs. 861-872. ISSN: 1694-609X. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1291147&lang=es&site=eds-live&scope=site>.
- [29] Alexander Bystritsky et al. «Current diagnosis and treatment of anxiety disorders». En: *Pharmacy and Therapeutics* 38.1 (2013), pág. 30.
- [30] Julia Elizabeth Calderón-Reyes et al. «Data Science Based Methodology: Design Process of a Correlation Model Between EEG Signals and Brain Regions Mapping in Anxiety». En: *New Perspectives in Software Engineering*. Ed. por Jezreel Mejia et al. Cham: Springer International Publishing, 2023, págs. 141-151. ISBN: 978-3-031-20322-0.
- [31] Julia Elizabeth Calderón-Reyes et al. «Methodology Design Of The Correlation Between EEG Signals And Brain Regions Mapping In Panic Attacks». En: *HCI International 2022 - Late Breaking Papers. Multimodality in Advanced Interaction Environments: 24th International Conference on Human-Computer Interaction, HCII 2022, Virtual Event, June 26 – July 1, 2022, Proceedings*. Berlin, Heidelberg: Springer-Verlag, 2022, págs. 357-370. ISBN: 978-3-031-17617-3. DOI: 10.1007/978-3-031-17618-0\_26. URL: [https://doi.org/10.1007/978-3-031-17618-0\\_26](https://doi.org/10.1007/978-3-031-17618-0_26).

- [32] Wenjun Cao et al. «The psychological impact of the COVID-19 epidemic on college students in China». En: *Psychiatry research* 287 (2020), pág. 112934.
- [33] Héctor Cardona-Reyes, Miguel Ángel Ortiz-Esparza y Jaime Muñoz-Arteaga. «Use of Learning Paths Through a Digital Ecosystem to Support Children With Learning Problems in Basic Math». En: *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje* 17.1 (2022), págs. 79-88. DOI: 10.1109/RITA.2022.3149784.
- [34] Jaramillo Carlos Mario Zapata, Giraldo Gloria L. y Giraldo Germán A. Urrego. «Las ontologías en la ingeniería de software: un acercamiento de dos grandes áreas del conocimiento Ontologies in software engineering: approaching two great knowledge areas.» En: *Revista Ingenierías Universidad de Medellín* 9.16 (2010), págs. 91-99. ISSN: 1692-3324. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsdoj&AN=edsdoj.1c6de204da04c1e88802a4a41adda11&lang=es&site=eds-live&scope=site>.
- [35] Chao Chen et al. «Efficacy Evaluation of Neurofeedback-Based Anxiety Relief». En: *Frontiers in Neuroscience* 15 (2021).
- [36] Alejandro Cifuentes-Muñoz. «Supuestos epistemológicos en corrientes psicoterapéuticas fundamentales: Una propuesta metodológica.» En: *Interdisciplinaria: Revista de Psicología y Ciencias Afines* 38.2 (2021), págs. 149-166. ISSN: 03258203. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=150600242&lang=es&site=eds-live&scope=site>.
- [37] Lorena Cudris-Torres, Álvaro Barrios-Núñez y Nidia J Bonilla-Cruz. «Coronavirus: epidemia emocional y social». En: (2020).
- [38] Pratap Dangeti. *Statistics for machine learning*. Packt Publishing Ltd, 2017.
- [39] Saloni Dattani, Hannah Ritchie y Max Roser. *Mental health*. [Online]. 2021. URL: <https://ourworldindata.org/mental-health>.
- [40] Geraldine Ed Dawson y Kurt W Fischer. *Human behavior and the developing brain*. The Guilford Press, 1994, págs. 3-11.
- [41] Clelia Manfredi De Poderoso, M. Cristina Julian y Leonardo Linetzky. «Reflexiones sobre trauma y repetición a partir de la clínica con pacientes con pánico y estrés posttraumático: caso Cromagnon». En: *Psicoanálisis: Revista de la Asociación Psicoanalítica de Buenos Aires* 29.1 (2007), págs. 165-179. ISSN: 0325819X. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/>

- login.aspx?direct=true&db=lth&AN=25813459&lang=es&site=eds-live&scope=site.
- [42] Amanda Díaz-García et al. «Efficacy of a Transdiagnostic internet-based treatment for emotional disorders with a specific component to address positive affect: Study protocol for a randomized controlled trial». En: *BMC psychiatry* 17.1 (2017), págs. 1-14.
- [43] DIF. *DIF Estatal Sistema para el Desarrollo Integral de la Familia*. [Online]. 2022. URL: <https://www.aguascalientes.gob.mx/dif/>.
- [44] Patrick Doherty et al. *Knowledge Representation Techniques. recurso electrónico : A Rough Set Approach*. Studies in Fuzziness and Soft Computing: 202. Springer, 2006. ISBN: 978-3-540-33518-4. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cat04898a&AN=ua.000157842&lang=es&site=eds-live&scope=site>.
- [45] EPOC EMOTIV. *Channel Mobile Brainwear| EMOTIV*. [Online]. 2022. URL: <https://www.emotiv.com/product/emotiv-epoc-x-14-channel-mobile-brainwear/>.
- [46] Murray Esler et al. «Human sympathetic nerve biology: parallel influences of stress and epigenetics in essential hypertension and panic disorder». En: *Annals of the New York Academy of Sciences* 1148.1 (2008), págs. 338-348.
- [47] M Faisal, E M Zamzami y Sutarman. «Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance.» En: *Journal of Physics: Conference Series; volume 1566, issue 1, page 012112; ISSN 1742-6588 1742-6596* (2020). URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.C4DE193D&lang=es&site=eds-live&scope=site>.
- [48] PM Fisher y AR Hariri. «Linking variability in brain chemistry and circuit function through multimodal human neuroimaging». En: *Genes, Brain and Behavior* 11.6 (2012), págs. 633-642.
- [49] «fMRI amygdala activation during a spontaneous panic attack in a patient with panic disorder». En: *The world journal of biological psychiatry: the official journal of the World Federation of Societies of Biological Psychiatry* 8 (4 2007), págs. 269-272. ISSN: 1562-2975. DOI: 10.1080/15622970701216673. URL: <https://pubmed.ncbi.nlm.nih.gov/17853295/>.

- [50] Jonathan Follett. *What is Lean UX?* [Online]. 2017. URL: <https://www.oreilly.com/radar/what-is-lean-ux/>.
- [51] Rita Francese, Michele Risi y Genoveffa Tortora. «A user-centered approach for detecting emotions with low-cost sensors». En: *Multimedia Tools and Applications* 79.47 (2020), págs. 35885-35907.
- [52] Walter J. Freeman. *Imaging Brain Function With EEG: Advanced Temporal and Spatial Analysis of Electroencephalographic Signals*. Springer New York, 2013. ISBN: 978-1-4614-4984-3. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edshlc&AN=edshlc.013716291.X&lang=es&site=eds-live&scope=site>.
- [53] Hamido Fujita y Guido Guizzi. *Intelligent Software Methodologies, Tools and Techniques: 14th International Conference, SoMet 2015, Naples, Italy, September 15-17, 2015. Proceedings*. Sep. de 2015. ISBN: 9783319226880.
- [54] Pérez George Argota. «Barreras epistemológicas para la arquitectura de los datos y la significación en el modelo predictivo de la ciencia.» En: *Campus* 25.30 (2020), págs. 329-335. ISSN: 1812-6049. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsdoj&AN=edsdoj.61bcd8ebfc0e4178a7a6a4856b3d712b&lang=es&site=eds-live&scope=site>.
- [55] Juliana Gonzalez-Astudillo et al. «Network-based brain-computer interfaces: principles and applications.» En: *Journal of neural engineering* 18.1 (2021). ISSN: 1741-2552. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cmedm&AN=33147577&lang=es&site=eds-live&scope=site>.
- [56] Jeff Gothelf y Josh Seiden. *Lean UX*. "O'Reilly Media, Inc.", 2021.
- [57] Bernhard Graimann, Gert Pfurtscheller y Brendan Allison. *Brain-Computer Interfaces. recurso electrónico : Revolutionizing Human-Computer Interaction*. The Frontiers Collection. Springer-Verlag Berlin Heidelberg, 2010. ISBN: 978-3-642-02090-2. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cat04898a&AN=ua.000170585&lang=es&site=eds-live&scope=site>.

- [58] Gillian Grennan et al. «Cognitive and Neural Correlates of Loneliness and Wisdom during Emotional Bias.» En: (2021). URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edssch&AN=edssch.oai%5C%3aescholarship.org%5C%2fark%5C%3a%5C%2f13030%5C%2fq42d670fp&lang=es&site=eds-live&scope=site>.
- [59] Rosa Mayelín Guerra Breña, Edgar Enrique Orozco Inca y María Beatriz Valencia Bonilla. «Las normas técnicas para la formación en ingeniería biomédica, tecnología y administración en salud.» En: *Revista Cubana de Educación Médica Superior* 36.3 (2022), págs. 1-18. ISSN: 08642141. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=lth&AN=159648975&lang=es&site=eds-live&scope=site>.
- [60] Tunç Gültekin y Aybars Uğur. «A hybrid approach based on transfer and ensemble learning for improving performances of deep learning models on small datasets.» En: *Turkish Journal of Electrical Engineering & Computer Sciences* 29.7 (2021), págs. 3197-3211. ISSN: 13000632. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=153990960&lang=es&site=eds-live&scope=site>.
- [61] J. A. Hartigan y M. A. Wong. «Algorithm AS 136: A K-Means Clustering Algorithm.» En: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1 (1979), págs. 100-108. ISSN: 00359254. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsjsr&AN=edsjsr.10.2307.2346830&lang=es&site=eds-live&scope=site>.
- [62] Hilary Jacobs Hendel. *Anxiety And Fear: What's The Difference?* [Online]. 2021. URL: <https://www.nami.org/Blogs/NAMI-Blog/May-2021/Anxiety-And-Fear-What-s-The-Difference>.
- [63] Te-Ming Huang, Vojislav Kecman e Ivica Kopriva. «Kernel Based Algorithms for Mining Huge Data Sets». En: *Kernel Based Algorithms for Mining Huge Data Sets* (2006). DOI: 10.1007/3-540-31689-2.
- [64] Casanova Hugo. «Estadística y Ciencia de Datos ¿Qué hay de nuevo?.» En: *Actualidad Económica* 32.107 (2022), págs. 55-75. ISSN: 2250-754X. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsdoj&AN=edsdoj.4b66653092c64e48afe1e748af6cc8a6&lang=es&site=eds-live&scope=site>.

- [65] David Hutchison et al. «Fundamentals of Inclusive HCI Design.» En: *Universal Access in Human Computer Interaction. Coping with Diversity*. 2007, págs. 3-9. ISBN: 9783540732785. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edo&AN=33420248&lang=es&site=eds-live&scope=site>.
- [66] «Increased Depression and Anxiety Disorders during the COVID-19 Pandemic in Children and Adolescents: A Literature Review». En: *Life (Basel, Switzerland)* 11 (11 nov. de 2021). ISSN: 2075-1729. DOI: 10.3390/LIFE11111188. URL: <https://pubmed.ncbi.nlm.nih.gov/34833064/>.
- [67] Iñaki Iturrate, Ricardo Chavarriaga y José del R. Millán. «General principles of machine learning for brain-computer interfacing.» En: (2020). ISSN: 978-0-444-63934-9. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.882F1336&lang=es&site=eds-live&scope=site>.
- [68] Asha P. Johnson et al. «Bioinspired and biomimetic micro- and nanostructures in biomedicine.» En: *Journal of Controlled Release* 343 (2022), págs. 724-754. ISSN: 0168-3659. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=S0168365922000888&lang=es&site=eds-live&scope=site>.
- [69] Rivero José Javier Leal. «Ciencia de datos e inteligencia artificial como apoyo para investigaciones cualitativas.» En: *Educare* 26.2 (2022). ISSN: 1316-6212. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsdoj&AN=edsdoj.9bb5cb3006ac49f7b2e1d01036185b74&lang=es&site=eds-live&scope=site>.
- [70] Olson Judith S. y Kellogg Wendy A. *Ways of Knowing in HCI*. Springer, 2014. ISBN: 9781493903771. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=2532841&lang=es&site=eds-live&scope=site>.
- [71] Eugenijus Kaniusas. *Fundamentals of Biosignals*. Springer Berlin Heidelberg, 2011. ISBN: 978-3-642-24842-9. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.21BFB73B&lang=es&site=eds-live&scope=site>.
- [72] Aleksandra Kawala-Sterniuk et al. «Summary of over Fifty Years with Brain-Computer Interfaces—A Review». En: *Brain Sciences* 11.1 (2021), pág. 43.

- [73] Raza Khalid y Dey Nilanjan. *Translational Bioinformatics in Healthcare and Medicine*. Advances in Ubiquitous Sensing Applications for Healthcare volume 13. Academic Press, 2021. ISBN: 9780323898249. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=2665146&lang=es&site=eds-live&scope=site>.
- [74] Souhir Khessiba et al. «Innovative deep learning models for EEG-based vigilance detection». En: *Neural Computing and Applications 2020 33:12 33* (12 nov. de 2020), págs. 6921-6937. ISSN: 1433-3058. DOI: 10.1007/S00521-020-05467-5. URL: <https://link.springer.com/article/10.1007/s00521-020-05467-5>.
- [75] Katharina Kircanski et al. «Subtypes of panic attacks: A critical review of the empirical literature». En: *Depression and Anxiety 26* (10 oct. de 2009), págs. 878-887. ISSN: 10914269. DOI: 10.1002/DA.20603.
- [76] Vitaly Kompaniets, Alexander Lyz y Alina Kazanskaya. «An Empirical Study of Goal Setting in UX/UI-design». En: *2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)*. 2020, págs. 1-5. DOI: 10.1109/AICT50176.2020.9368570.
- [77] Andrea Kübler. «The history of BCI: From a vision for the future to real support for personhood in people with locked-in syndrome». En: *Neuroethics 13.2* (2020), págs. 163-180.
- [78] Masaaki Kurosu. *Human-Computer Interaction. Theories, Methods, and Tools: 16th International Conference, HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I*. Lecture Notes in Computer Science. 2014. ISBN: 978-3-319-07233-3. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edshlc&AN=edshlc.014106178.2&lang=es&site=eds-live&scope=site>.
- [79] Sebastian Kurtek et al. «Segmentation, alignment and statistical analysis of biosignals with application to disease classification.» En: *Journal of Applied Statistics 40.6* (2013), págs. 1270-1288. ISSN: 02664763. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=87479371&lang=es&site=eds-live&scope=site>.
- [80] Paul Kvam, Brani Vidakovic y Seong-joon Kim. *Nonparametric Statistics with Applications to Science and Engineering with R*. Vol. 1. John Wiley & Sons, 2022.

- [81] Gary W. Ladd. «Peer Rejection, Aggressive or Withdrawn Behavior, and Psychological Maladjustment from Ages 5 to 12: An Examination of Four Predictive Models.» En: *Child Development* 77.4 (2006), págs. 822-846. ISSN: 00093920. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsjsr&AN=edsjsr.3878401&lang=es&site=eds-live&scope=site>.
- [82] Daniel Lemus-Delgado y Ricardo Pérez Navarro. «Ciencia de datos y estudios globales: aportaciones y desafíos metodológicos.» En: *Colombia Internacional* 102 (2020), págs. 41-62. ISSN: 01215612. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=143018624&lang=es&site=eds-live&scope=site>.
- [83] R. Leote. «Biomimicry and Art: Transductions with Biology.» En: *Revista de Humanidades de Valparaíso* 18 (2021), págs. 225-243 -243. ISSN: 07194242. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-85130053441&lang=es&site=eds-live&scope=site>.
- [84] Zina Li, Shuqing Zhang y Jiahui Pan. «Advances in Hybrid Brain-Computer Interfaces: Principles, Design, and Applications.» En: *Computational intelligence and neuroscience* 2019 (2019), pág. 3807670. ISSN: 1687-5273. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=31687006&lang=es&site=eds-live&scope=site>.
- [85] Nan Lin, Douglas Noe y Xuming He. «Tree-Based Methods and Their Applications.» En: *Springer Handbooks* (2006), págs. 551-570. ISSN: 25228706. DOI: 10.1007/978-1-84628-288-1\_30/TABLES/5. URL: [https://link-springer-com.dibpxy.uaa.mx/referenceworkentry/10.1007/978-1-84628-288-1\\_30](https://link-springer-com.dibpxy.uaa.mx/referenceworkentry/10.1007/978-1-84628-288-1_30).
- [86] Fan Liu y Yong Deng. «Determine the Number of Unknown Targets in Open World Based on Elbow Method.» En: *IEEE Transactions on Fuzzy Systems* 29.5 (2021), págs. 986-995. DOI: 10.1109/TFUZZ.2020.2966182.
- [87] Clara Isabel López, Luis Eduardo Bautista y Carlos Mantilla. «A Novel Approach to Learning Virtual Engineering in the Medical Devices Process.» En: *Education and Information Technologies* 25.6 (2020), págs. 5703-5729. ISSN: 1360-2357. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1276187&lang=es&site=eds-live&scope=site>.

- [88] Huizilopoztli Luna-García, Francisco Javier Álvarez-Rodríguez y Ricardo Mendoza-González. «Modelo de gestión para diseño curricular basado en prácticas de ingeniería de software.» En: *Revista Electrónica de Investigación Educativa* 17.3 (2015), págs. 61-78. ISSN: 16074041. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=109369775&lang=es&site=eds-live&scope=site>.
- [89] Chanaka Mannapperuma. «Human-computer interaction principles for developing web-based genomics resources.» En: (2020). ISSN: 91-7855-284-2. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsswe&AN=edsswe.oai.DIVA.org.umu.170109&lang=es&site=eds-live&scope=site>.
- [90] Antonio Costantino Marceddu et al. «A Novel Redundant Validation IoT System for Affective Learning Based on Facial Expressions and Biological Signals.» En: *Sensors* 22.7 (2022). ISSN: 14248220. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-85127589711&lang=es&site=eds-live&scope=site>.
- [91] Iñigo Martínez, Elisabeth Viles e Igor G. Olaizola. «Data Science Methodologies: Current Challenges and Future Approaches». En: *Big Data Research*. 24 (mayo de 2021), pág. 100183. ISSN: 2214-5796. DOI: 10.1016/J.BDR.2020.100183.
- [92] Ana Masedo et al. «A multicentric study on stigma towards people with mental illness in health sciences students». En: *BMC Medical Education* 21.1 (2021), págs. 1-11.
- [93] Juliano Paulo Menzen, Kleinner Farias y Vinicius Bischoff. «Using biometric data in software engineering: a systematic mapping study.» En: *Behaviour & Information Technology* 40.9 (2021), págs. 880-902. ISSN: 0144929X. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=152059714&lang=es&site=eds-live&scope=site>.
- [94] Evangelia Micheli-Tzanakou et al. «Biosignals and Compression Standards.» En: *M-Health*. 2006, págs. 277-292. ISBN: 9780387265582. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edo&AN=32839695&lang=es&site=eds-live&scope=site>.
- [95] Prabhaker Mishra et al. «Scales of measurement and presentation of statistical data». En: *Annals of cardiac anaesthesia* 21.4 (2018), pág. 419.

- [96] Hugo A. Mitre-Hernández et al. «Strategies for fostering Knowledge Management Programs in Public Organizations.» En: (2015). URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.357FD26A&lang=es&site=eds-live&scope=site>.
- [97] Daniel B Morehead. *Science Over Stigma: Education and Advocacy for Mental Health*. American Psychiatric Pub, 2021.
- [98] Farah Muhammad y Saad Al-Ahmadi. «Human state anxiety classification framework using EEG signals in response to exposure therapy.» En: *PLoS ONE* 17.3 (2022), e0265679. ISSN: 1932-6203. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsgea&AN=edsgcl.697334451&lang=es&site=eds-live&scope=site>.
- [99] José Muñiz Fernández et al. «Tratamientos Psicológicos Empíricamente Apoyados Para Adultos: Una Revisión Selectiva.» En: *Psicothema* 33.2 (2021), págs. 188-197. ISSN: 02149915. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=149980303&lang=es&site=eds-live&scope=site>.
- [100] Sergio Ignacio Muñoz-Fernández et al. «Estrés, respuestas emocionales, factores de riesgo, psicopatología y manejo del personal de salud durante la pandemia por COVID-19». En: *Acta Pediátrica de México* 41.S1 (2020), págs. 127-136.
- [101] Ayman Nadeem. «Human-centered approach to static-Analysis-driven developer tools». En: *Communications of the ACM* 65 (3 mar. de 2022), págs. 38-45. ISSN: 15577317. DOI: 10.1145/3486597.
- [102] Felix Naumann. «Data profiling revisited». En: *ACM SIGMOD Record* 42.4 (2014), págs. 40-49.
- [103] Amy Novotney. «Continuing education the risks of social isolation». En: *Monitor on Psychology* 50.5 (2019), págs. 32-37.
- [104] Abiodun Afolayan Ogunyemi et al. «A Systematic Mapping Study of HCI Practice Research.» En: *International Journal of Human-Computer Interaction* 35.16 (2019), págs. 1461-1486. ISSN: 10447318. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=138051558&lang=es&site=eds-live&scope=site>.
- [105] World Health Organization. *Mental health*. URL: <https://www.who.int/en/news-room/fact-sheets/detail/mental-health-strengthening-our-response>.

- [106] Pedro García Parajuá y Mónica Magaritos López. *Guía de Bolsillo de la Clasificación CIE-10: Clasificación de los Trastornos Mentales y del Comportamiento*. Ed. Médica Panamericana, 2000.
- [107] K. Peng et al. «Named Entity Aware Transfer Learning for Biomedical Factoid Question Answering.» En: *IEEE/ACM Transactions on Computational Biology and Bioinformatics, Computational Biology and Bioinformatics, IEEE/ACM Transactions on, IEEE/ACM Trans. Comput. Biol. and Bioinf* 19.4 (2022), págs. 2365-2376. ISSN: 1545-5963. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsee&AN=edsee.9428547&lang=es&site=eds-live&scope=site>.
- [108] Sandra GL Persiani et al. «Biometric data as real-time measure of physiological reactions to environmental stimuli in the built environment.» En: *Energies* 14.1 (2021). ISSN: 19961073. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-85106150679&lang=es&site=eds-live&scope=site>.
- [109] Jonas Peters, Dominik Janzing y Bernhard Schölkopf. «Elements of Causal Inference». En: *Foundations and Learning Algorithms* (2017), pág. 288.
- [110] Mónica Pieschacón Fonrodona. *Trastorno de estrés postraumático. Modelo cognitivo-conductual*. Ediciones Uniandes, 2011. ISBN: 9789586956598. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=1630483&lang=es&site=eds-live&scope=site>.
- [111] Kevin M. Pitt et al. «Behind the Scenes of Noninvasive Brain-Computer Interfaces: A Review of Electroencephalography Signals, How They Are Recorded, and Why They Matter.» En: *Perspectives of the ASHA Special Interest Groups* 4.6 (2019), págs. 1622-1636. ISSN: 2381-473X. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ccm&AN=141143749&lang=es&site=eds-live&scope=site>.
- [112] J. Quinlan. «Induction of decision trees.» En: *Machine Learning* 1.1 (1986), págs. 81-106. ISSN: 08856125. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=71081842&lang=es&site=eds-live&scope=site>.

- [113] J. R. (John Ross) Quinlan. *C4.5: programs for machine learning*. The Morgan Kaufmann series in machine learning. Morgan Kaufmann Publishers, 1993. ISBN: 978-1-55860-238-0. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edshlc&AN=edshlc.005934215.3&lang=es&site=eds-live&scope=site>.
- [114] Martina Ferández Raone y Sergio Rubén Zanassi. «Antecedentes y revisión crítica del denominado Ataque de pánico». En: *Perspectivas en Psicología: Revista de Psicología y Ciencias Afines* 13.2 (2016), págs. 57-66.
- [115] Jens F Rehfeld. «Cholecystokinin and Panic Disorder: Reflections on the History and Some Unsolved Questions». En: *Molecules* 26.18 (2021), pág. 5657.
- [116] Juan José Rey-Serrano y Miguel Enrique Ochoa-Vera. «La enseñanza de la bioética para el fortalecimiento de la práctica biomédica y la investigación en las escuelas de ciencias de la salud.» En: *MedUNAB* 25.2 (2022), págs. 246-252. ISSN: 01237047. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=lth&AN=159512447&lang=es&site=eds-live&scope=site>.
- [117] David Riaño et al. *Knowledge Representation for Health-Care. recurso electrónico : ECAI 2010 Workshop KR4HC 2010, Lisbon, Portugal, August 17, 2010, Revised Selected Papers*. Lecture Notes in Computer Science: 6512. Springer Berlin Heidelberg, 2011. ISBN: 978-3-642-18049-1. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cat04898a&AN=ua.000176043&lang=es&site=eds-live&scope=site>.
- [118] John B. Rollins. «Metodología Fundamental para la Ciencia de Datos». En: (2015).
- [119] Wolfram Schultz. «Dopamine signals for reward value and risk: basic and recent data». En: *Behavioral and brain functions* 6.1 (2010), págs. 1-9.
- [120] Claude E Shannon. *Claude Elwood Shannon: Collected papers*. IEEE press, 1993.
- [121] Yunzhi Shao. «Human-Computer Interaction Environment Monitoring and Collaborative Translation Mode Exploration Using Artificial Intelligence Technology.» En: *Journal of environmental and public health* 2022 (2022), pág. 4702003. ISSN: 1687-9813. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=36213014&lang=es&site=eds-live&scope=site>.

- [122] Zeguo Shao et al. «Influences of Daily Life Habits on Risk Factors of Stroke Based on Decision Tree and Correlation Matrix». En: *Computational and mathematical methods in medicine 2020* (2020). ISSN: 1748-6718. DOI: 10.1155/2020/3217356. URL: <https://pubmed.ncbi.nlm.nih.gov/32565878/>.
- [123] Shivani Sharma y Durga Toshniwal. «MR-OVnTSA: a heuristics based sensitive pattern hiding approach for big data». En: *Applied Intelligence* 50 (12 dic. de 2020), págs. 4241-4260. ISSN: 15737497. DOI: 10.1007/S10489-020-01749-6/FIGURES/6. URL: <https://link.springer.com/article/10.1007/s10489-020-01749-6>.
- [124] Kamila Słupińska, Jarosław Duda y Konrad Biercewicz. «Planning an experiment in a virtual environment reality as a place of research on human behaviour using methods of neuroscience measurement – bibliometric analysis and methodological approach.» En: *Procedia Computer Science* 192 (2021), págs. 3123-3133. ISSN: 18770509. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edo&AN=152766904&lang=es&site=eds-live&scope=site>.
- [125] Tony C. Smith y Eibe Frank. «Introducing Machine Learning Concepts with WEKA.» En: *Statistical Genomics* (2016), págs. 353-378. ISSN: 9781493935765. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=116047126&lang=es&site=eds-live&scope=site>.
- [126] Guillermo Soberón-Acevedo. «I. Estructura y funciones de la Secretaría de Salud, del Sector Salud y del Sistema Nacional de Salud». En: *Salud Pública de México* 29.2 (1987), págs. 127-140.
- [127] Posit Software. *Download RStudio*. 2023. URL: <https://posit.co/downloads/>.
- [128] Stefan Studer et al. «Towards CRISP-ML (Q): a machine learning process model with quality assurance methodology». En: *Machine learning and knowledge extraction* 3.2 (2021), págs. 392-413.
- [129] Nazmi Sofian Suhaimi, James Mountstephens y Jason Teo. «EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities». En: *Computational intelligence and neuroscience 2020* (2020). ISSN: 1687-5273. DOI: 10.1155/2020/8875426. URL: <https://pubmed.ncbi.nlm.nih.gov/33014031/>.

- [130] Huaining Sun, Xuegang Hu y Yuhong Zhang. «Attribute Selection Based on Constraint Gain and Depth Optimal for a Decision Tree». En: *Entropy (Basel, Switzerland)* 21 (2 feb. de 2019). ISSN: 1099-4300. DOI: 10 . 3390 / E21020198. URL: <https://pubmed.ncbi.nlm.nih.gov/33266913/>.
- [131] Desney S. Tan y Anton Nijholt. *Brain-Computer Interfaces. recurso electrónico : Applying our Minds to Human-Computer Interaction*. Human-Computer Interaction Series. Springer-Verlag London Limited, 2010. ISBN: 978-1-84996-271-1. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=cat04898a&AN=ua.000171640&lang=es&site=eds-live&scope=site>.
- [132] R Core Team. *R: A Language and Environment for Statistical Computing*. 2023. URL: <https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf>.
- [133] Anja Thieme, Danielle Belgrave y Gavin Doherty. «Machine Learning in Mental Health: A Systematic Review of the HCI Literature to Support the Development of Effective and Implementable ML Systems.» En: *ACM Transactions on Computer-Human Interaction (TOCHI)* 27.5 (2020), págs. 1-53. ISSN: 10730516. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=aci&AN=145359360&lang=es&site=eds-live&scope=site>.
- [134] Tim Thornton. *Psychiatric diagnosis, tacit knowledge, and criteria*. Oxford University Press, 2016. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.8911916B&lang=es&site=eds-live&scope=site>.
- [135] Umesh Kumar Tiwari y Santosh Kumar. *Component-based software engineering: Methods and metrics*. CRC Press, 2020.
- [136] Arshya Vahabzadeh, Charles F Gillespie y Kerry J Ressler. «Fear-related anxiety disorders and post-traumatic stress disorder». En: *Neurobiology of Brain Disorders* (2015), págs. 612-620.
- [137] Peter Van Schuerbeek et al. «Does the amygdala response correlate with the personality trait ‘harm avoidance’ while evaluating emotional stimuli explicitly?» En: *Behavioral and Brain Functions* 10.1 (2014), págs. 1-13.

- [138] Bajaj Varun y Sinha G. R. *Artificial Intelligence-Based Brain-Computer Interface*. Academic Press, 2022. ISBN: 9780323911979. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=3060343&lang=es&site=eds-live&scope=site>.
- [139] *What are the five major types of anxiety disorders?* | HHS.gov. URL: <https://www.hhs.gov/answers/mental-health-and-substance-abuse/what-are-the-five-major-types-of-anxiety-disorders/index.html>.
- [140] *What's the difference between stress and anxiety?* [Online]. 2019. URL: <https://www.apa.org/topics/stress/anxiety-difference>.
- [141] Adam B Wilcox, David K Vawdrey y Kensaku Kawamoto. *Software engineering for health care and biomedicine*. Springer London, 2014. ISBN: 9781447144748. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-85028513853&lang=es&site=eds-live&scope=site>.
- [142] Jonathan R. Wolpaw, José del R. Millán y Nick F. Ramsey. «Chapter 2 - Brain-computer interfaces: Definitions and principles.» En: *Handbook of Clinical Neurology* 168 (2020), págs. 15-23. ISSN: 978-0-444-63934-9. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edselp&AN=B9780444639349000020&lang=es&site=eds-live&scope=site>.
- [143] Yunfeng Wu y Behnaz Ghoraani. «Biological Signal Processing and Analysis for Healthcare Monitoring.» En: *Sensors (Basel, Switzerland)* 22.14 (2022). ISSN: 1424-8220. URL: <https://dibpxy.uaa.mx/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=mdc&AN=35891021&lang=es&site=eds-live&scope=site>.
- [144] Shuo Yang, Jing Zhi Guo y Jun Wei Jin. «An improved Id3 algorithm for medical data classification». En: *Computers & Electrical Engineering* 65 (ene. de 2018), págs. 474-487. ISSN: 0045-7906. DOI: 10.1016/J.COMPELECENG.2017.08.005.
- [145] Adrián Yoris et al. «The roles of interoceptive sensitivity and metacognitive interoception in panic». En: *Behavioral and Brain Functions* 11.1 (2015), págs. 1-6.
- [146] Haojie Zhang et al. «Fine-Tuning Pre-Trained Language Models Effectively by Optimizing Subnetworks Adaptively.» En: (2022). URL: <https://dibpxy.uaa.mx/>

login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsarx&AN=edsarx.2211.01642&lang=es&site=eds-live&scope=site.

- [147] Joe Zhangid et al. «Best practices in the real-world data life cycle». En: *PLOS Digital Health* 1 (1 ene. de 2022), e0000003. ISSN: 2767-3170. DOI: 10.1371/JOURNAL.PDIG.0000003. URL: <https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000003>.
- [148] Lori A Zoellner y Edna B Foa. «Applying Research Domain Criteria (RDoC) to the study of fear and anxiety: A critical comment». En: *Psychophysiology* 53.3 (2016), págs. 332-335.





## A. Iteraciones de la Metodología

Las etapas de la propuesta metodológica fueron divididas acorde a los procesos de las subetapas que las integran, ello para la designación de *MVPs* y monitoreo de las iteraciones de la metodología bajo el esquema de:

- **Diseño Arquitectónico del Modelo.** Conjunto de etapas base de la metodología que dan seguimiento al planteamiento del problema de investigación, mediante la redefinición de preguntas de investigación, objetivos, y requerimientos; consenso entre el equipo multidisciplinario, las necesidades del usuario, e interacciones de los *stakeholders* para el bosquejo y reestructuración de componentes.
  - Segmentación: etapa de apertura de la metodología que permitió conformar el análisis de requerimientos mediante la definición de la población muestra, conjuntando la identificación de usuarios objetivo con el desarrollo de la solución a la problemática planteada en la transición a cada una de las etapas; formulación de objetivos e intervalos de los *MVPs*, creando puntos de referencia.
  - Diseño: planteamiento del esquema, información con base a la segmentación de datos en crudo provista por el Muestreo; modelo para las proyecciones de resultados de los *MVPs* en función de los niveles de interacción del usuario, requerimientos identificados en un análisis preliminar, y representación esquemática de los procesos y tareas principales del flujo de trabajo.
- **Dispersión y Asignación de Procesos.** Caracterización de resultados parciales de las iteraciones de la metodología, a través de directrices para la proyección de hallazgos correspondientes a los *MVPs*, análisis de requerimientos y esquema de diseño; designación del flujo de datos acorde al desglose de los resultados provistos, distribución de la información y tendencias asociadas a los procesos de las subetapas.
  - Tratamiento: especificaciones de la estructura, tipos de datos y variables a ser representadas como subconjunto de muestra a ser empleado a manera de prueba en la etapa de lanzamiento como el modelo de datos para la ejecución y evaluación de técnicas computacionales.
  - Lanzamiento: implementación de técnicas de IA aptas para los subconjuntos de datos, restricciones de costo y tiempo inherentes a la problemática; clasificación temprana de la solución implementada a manera de prueba mediante la aplicación del algoritmo de *K-Means* para la agrupación e *ID3* y *J48* para la clasificación.

- Evaluación: procesos para la verificación y validación del modelo sobre resultados parciales y porcentaje de precisión validado mediante el *feedback* de los *stakeholders*, en que un miembro del equipo en conjunto con un especialista de la salud consideraron el diseño experimental del modelo como exitoso en la exploración de la caracterización de los datos y el comportamiento del modelo dadas las técnicas aplicadas para la clasificación y agrupación de los datos.

Fungiendo como directrices para la vinculación de proyecciones de los requerimientos a cubrir mediante la generación de *MVPs*, análisis de resultados obtenidos, reuniones del equipo multidisciplinario, retroalimentación, y procesos del ciclo de mejora continua. Acorde al modelo educativo de la **UAA**, y en seguimiento a los reportes de avance de **Conacyt**, se efectuaron cuatro iteraciones de la propuesta metodológica, las cuales se muestran a detalle en la Figura A.1 mediante el desglose de sus *MVPs*:



Figura A.1: Desglose de Iteraciones de la Metodología.

## B. Publicaciones Realizadas

El epicentro de la propuesta metodológica se suscitó mediante la participación en el Congreso *HCI International 2022*, puesto que con la retroalimentación obtenida de la propuesta se afinaron detalles de las etapas de la metodología, así como vertientes relacionadas; por tanto, la Figura B.1 muestra el certificado obtenido tras la presentación de la primera publicación realizada en torno a la investigación, bajo el título de *Methodology Design of the Correlation between EEG Signals and Brain Regions Mapping in Panic Attacks* [31]:



Figura B.1: Constancia de Presentación por Artículo en el *HCI 2022*.

Consecuente al desarrollo de la metodología orientada a Ciencia de Datos y el enfoque hacia el usuario, la presentación del artículo titulado *Data Science based Methodology: Design Process of a Correlation Model between EEG Signals and Brain Regions Mapping in Anxiety* [30] en el Congreso Internacional *CIMPS 2022* que se muestra en la Figura B.2 contribuyó a la integración de etapas de la metodología e identificación de avances en los *MVPs* mediante el diseño instrumental del modelo para técnicas algorítmicas de IA.

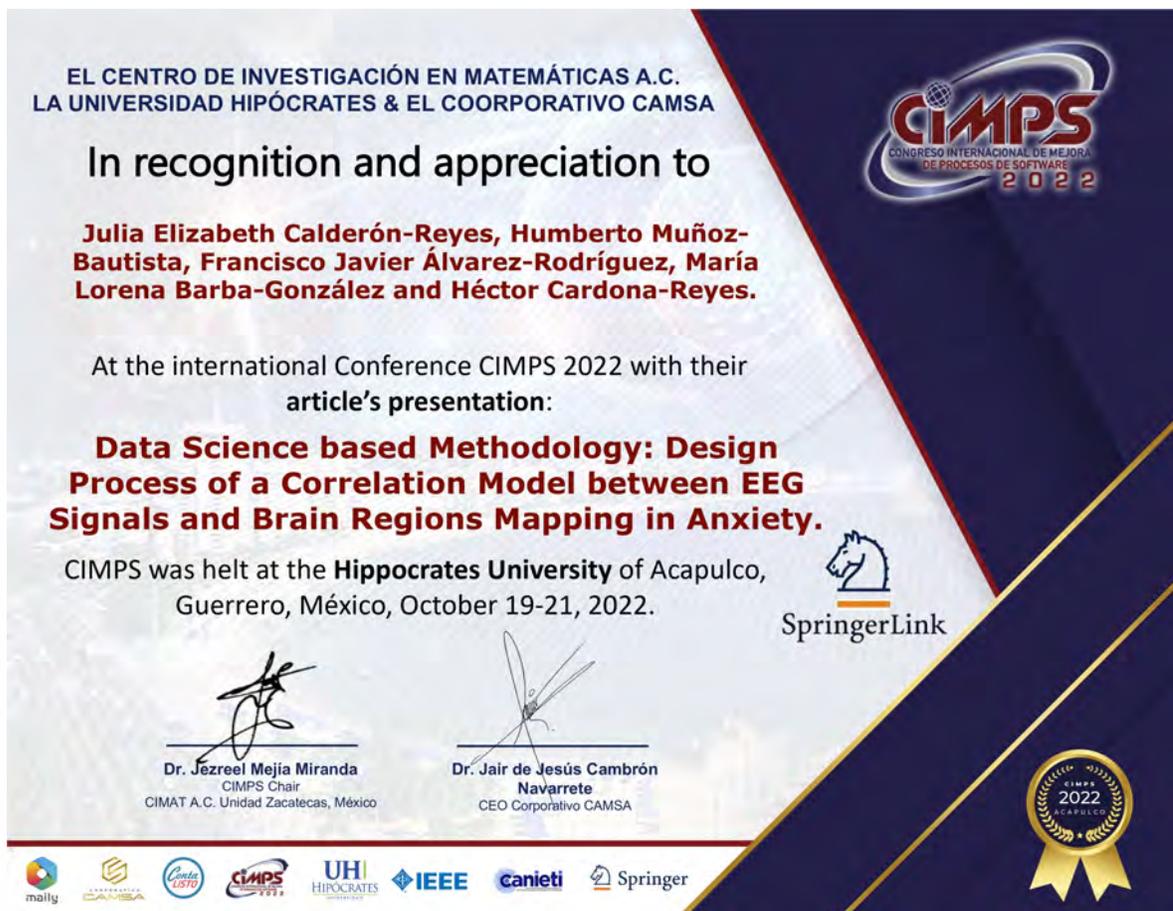


Figura B.2: Constancia de Presentación de Artículo en el *CIMPS 2022*.

Por otra parte, la participación en el taller de *Hybrid Intelligent Systems* en el marco del Congreso Internacional del *MICAI 2022* con el artículo titulado *Experimental Design of the Correlation Model between EEG Signals and Brain Regions Mapping in Anxiety* fue un punto culminante de la investigación realizada hasta el momento, misma que obtuvo reconocimiento como primer lugar de los artículos presentados bajo la modalidad de taller, tal como se muestra en la Figura B.3:



Figura B.3: Constancia de Primer Lugar en el MICA I 2022.

Consolidando las bases metodológicas mediante el análisis exploratorio del modelo y el diseño instrumental de este para la validación y verificación de los avances en torno al conjunto de datos analizados y los productos obtenidos hasta el momento.

## C. Trabajos a Futuro

Una vez realizado el análisis de los resultados, se procedió a realizar un esquema de trabajos a futuro, mismo que se muestra en la Figura C.1:

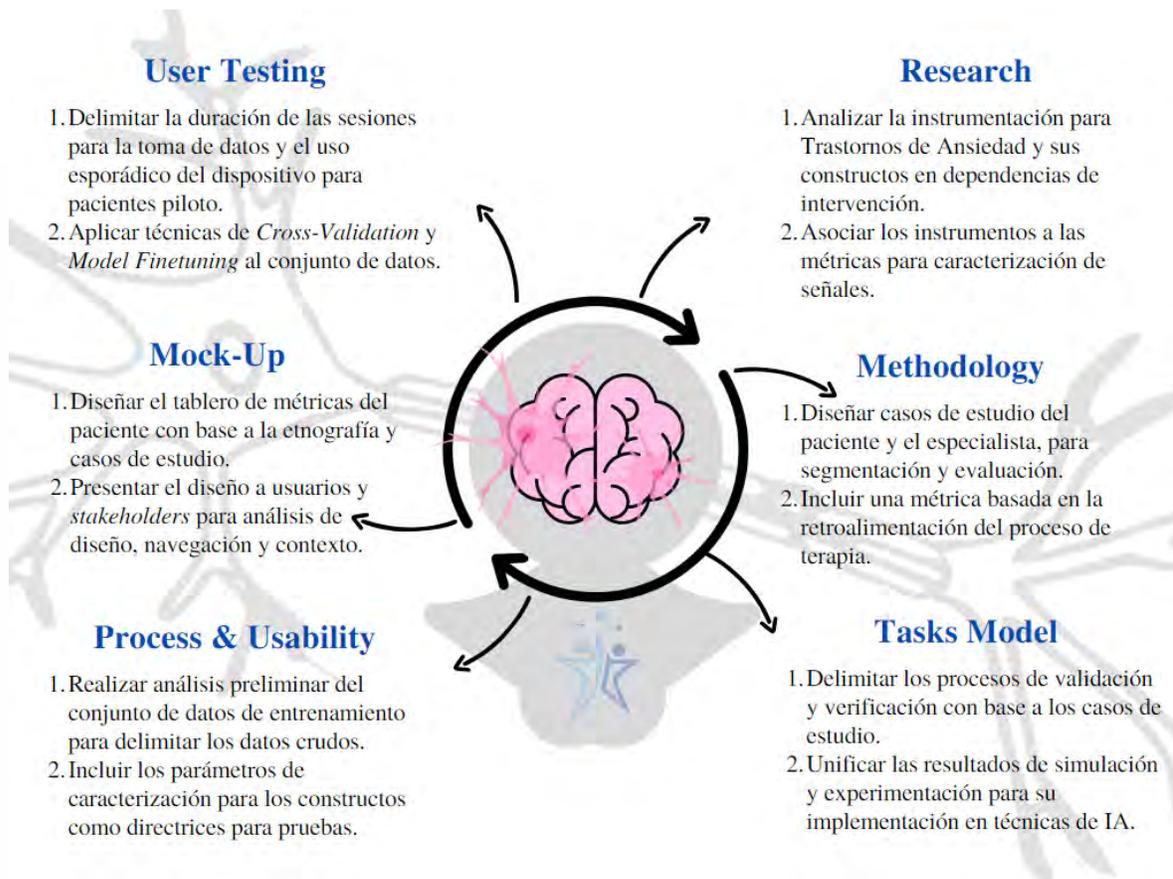


Figura C.1: Proyección de Trabajos a Futuro.

Dando prioridad a las pruebas de usuario e investigación continua para la adaptabilidad de la metodología como *Mock-Up*, sus procesos y usabilidad, así como modelo de tareas de forma puntual para el ciclo de mejora continua de la metodología, así como enriquecimiento del modelo; cabe mencionar, se identificaron vertientes de aplicación exponencial en bioinformática mediante la contribución a biobancos de datos para el estudio de trastornos neurológicos con la toma, monitoreo y recolección de bioseñales *EEG*, destacando así la importancia de la continuidad del trabajo de investigación.