



**INSTITUTO POLITÉCNICO
NACIONAL**



**UNIDAD PROFESIONAL INTERDISCIPLINARIA EN INGENIERÍA Y
TECNOLOGÍAS AVANZADAS**

Trabajo Terminal II

**“Sistema de monitoreo no invasivo del nivel de estrés a partir de
señales EEG y ECG”**

Que para obtener el título de
“Ingeniero en Biónica”

Presenta:

Jaime Hernández Ian Felipe

Asesores:

Dra. Blanca Tovar Corona

Dr. Adrián Octavio Ramírez Morales

Dr. Erick Dasaev de la Rosa Montero

Junio 2022

Índice general

Capítulo 1 Introducción	11
Capítulo 1.1 Estado del arte	12
Capítulo 2 Objetivo	18
2.1 Objetivo general	18
2.2 Objetivo específico	18
Capítulo 3 Planteamiento del problema	19
Capítulo 4 Justificación	20
Capítulo 5 Marco Teórico	21
5.1 Señal EEG	21
5.1.1 Potencial en neuronas	21
5.1.2 Sistema de conducción EEG	23
5.2 Electrocardiograma y señal ECG	23
5.2.1 Ancho de banda características ECG	24
5.2.2 Sistema de conducción cardiaco	24
5.3 La Variabilidad de la Frecuencia Cardiaca (VFC)	25
5.4 Estimación de nivel de estrés	25
5.4.1 Sistemas de instrumentación para filtrado de señales analógicas	26
5.5 Redes neuronales	26
5.6 Modelos de regresión multivariable	26
5.7 Árbol de decisión	27
5.8 Clúster	27
5.9 Análisis por factorial	27
5.10 Red neuronal Autoencoder	28
5.11 La Escala de Estrés Percibido (EEP)	28
Capítulo 6 Desarrollo experimental	29
6.1 Sujetos de estudio y tamaño de la muestra	29
6.2 Instrumentos	31
6.2.1 Mindflex	31
6.2.2 Sensor AD8232	32
6.3 Procedimientos:	33
6.3.1 Acondicionamiento de sensores ECG y EEG	17
6.4 Análisis teórico:	19

6.4.1 Descripción de variables análisis EEG	19
6.4.2 Preprocesamiento de datos y adquisición de datos	20
6.4.3 Correlación entre variables	20
6.4.4 Descripción de variables análisis ECG	23
6.4.5 Técnica no supervisada clasificación	27
6.4.6 Técnica supervisada algoritmo de clasificación	28
6.4.7 Muestreo EEP 10	30
6.4.8 Diseño de base de datos SQL	31
6.4.9 Diseño de esquema general del programa	33
Capítulo 7 Resultados	35
7.1 Análisis de Resultados	36
7.2 Interpretación de Resultados	37
Capítulo 7 Discusión	39
Capítulo 8 Conclusiones	40
Capítulo 10 Recomendaciones	41
Bibliografía	42
Anexo	46

Índice figuras

Figura 1 Sinapsis. Recuperado [28]	22
Figura 2 Representación de ancho de banda ECG. Recuperado [29]	24
Figura 3 Sistema de conducción cardíaco. Recuperado [30]	25
Figura 4 Diagrama de bloques solución propuesta.	33
Figura 5 Diagrama conexión Arduino UNO y sensor ECG. Recuperado [31]	17
Figura 6 Diagrama Arduino y Mindflex. Recuperado [29][32][33]	18
Figura 7 Diagrama de caja EEG	19
Figura 9 Análisis factorial EEG.	21
Figura 10 Radio de varianza explicada	22
Figura 11 Muestra de señales EEG (alta alfa, baja alfa y delta)	22
Figura 12 Muestra de señales ECG	24
Figura 13 Análisis de intervalos y distribución RR	25
Figura 14 Diagrama de árbol de decisión ECG	25
Figura 15 Base de datos con los seis predictores más importantes representados en porcentaje ECG.	26
Figura 16 Clúster análisis señal EEG	27
Figura 17 Relación de precisión y exhaustividad	30
Figura 18 Aplicación EPP 10 presentación resultados	31
Figura 19 Diseño base de datos SQL	32
Figura 20 Diagrama de flujo programa general	33
Figura 21 Presentación Power BI reporte global	36
Figura 22 Reporte de elementos personalizados.	38

Índice tablas

Tabla 1 Estudios previos estrés EEG. Modificado de [8]	12
Tabla 2 Muestras [24][25][26] [27] de proyectos de investigación y muestras adquiridas por los dispositivos y datos de prueba adquisición propia	30
Tabla 3 Correlaciones EEG	20
Tabla 4 Correlación r y absoluta	20
Tabla 5 Compendio de base de datos recopilada ECG [24]	23
Tabla 6 Base de datos con los seis predictores más importantes	24
Tabla 7 Base de datos con los seis predictores más importantes representados en porcentaje antes del pruning ECG.	26
Tabla 8 Resultado de análisis supervisado	28
Tabla 9 Red neuronal autoencoder	29
Tabla 10 Aplicación EPP 10 muestra 20 personas	31
Tabla 11 Características dispositivo EEG	35
Tabla 12 Características clúster analizando existencia de estrés en las señales ECG y EEG	38

Abreviaturas

CCN: Red neuronal convolucional
ECG: Electrocardiograma
EEG: Electroencefalograma
HF: Alta frecuencia en la variabilidad cardiaca
IA: Inteligencia Artificial
KNN: K vecinos más próximos
LDA: Análisis discriminante lineal
LF: Baja frecuencia en la variabilidad cardiaca
ML: Aprendizaje automático
NB: Clasificador bayesiano
NN: Red neuronal
NoSQL: Sistema de Gestión de Bases de Datos No Relacionales
PFC: Corteza prefrontal
RF: Clasificador bosque aleatorio
SQL: Sistema de Gestión de Bases de Datos Relacionales
SVM: Máquinas de vectores de soporte
VFC: Variabilidad de frecuencia cardiaca

Resumen

Este documento presenta un análisis estadístico, clasificador y extractor de características de las señales EEG y ECG, implementando un sistema de monitoreo a través de algoritmos de aprendizaje automático para su categorización. Gracias a este proceso se generó una segmentación del nivel de estrés que se encuentra sometida la persona. Con una muestra de 20 personas siendo nuestra población estadística, se adjuntó un mínimo de mil datos representativos de estímulos para generar la base de datos. También se usó una base de datos del MIT aumentando la muestra estadística a una n mayor de 110 personas, permitiendo etiquetar datos que ayudaron a realizar un almacenamiento de datos SQL y permite observar datos de interés. Se obtuvo un sistema de monitoreo que permite la visualización de la tendencia y un modelo matemático que describe la frecuencia a la que fue sometida una persona al aplicarle un estímulo que le ocasiona estrés.

Palabras claves: EEG, ECG, IA, ML, SQL, python, estrés, análisis y estadística

Abstract

This document consists of the statistical analysis of the EEG and ECG signals. It is a stress monitor system that will work with machine learning classifier algorithms. For this reason, it will be able to make a stress categorization for each person. Our sample is approximately twenty people, and it was our statistical values with a minimum of one thousand representative data of stress stimulations for making our database. And for testing our model we complemented with the MIT database, finally our database was an n of 110 people. It was a good option because we can use SQL databases for storing data. Then we used a statistical model for describing and visualizing trends and frequency of stress stimulations.

Key words: EEG, ECG, IA, ML, data, SQL, python, stress, analyzing and statistic

Glosario

Algoritmo: El núcleo de donde parte la inteligencia artificial. Son las fórmulas matemáticas y/o comandos de programación que indican a una computadora cómo resolver problemas. Son las reglas para enseñar a las computadoras cómo resolver situaciones o problemas.

Aprendizaje no supervisado: En el modelo de Inteligencia Artificial por sus siglas IA se proporciona la respuesta correcta con anticipación: la IA conoce tanto la pregunta como la respuesta. Este método de preparación es el más común, porque define los modelos de pregunta y respuesta ofreciendo la mayor cantidad de datos.

Aprendizaje supervisado: Los modelos de IA pueden aprender por sí mismos, sin tener que alimentarlos con estructuras predefinidas. Utilizan capas y capas de información no estructurada, procesan los datos, establecen las relaciones existentes entre ellos y encuentran un patrón en los mismos.

Árbol de decisión: Es un modelo predictivo que divide el espacio de los predictores agrupando observaciones con valores similares para la variable respuesta o dependiente.

Correlaciones: Cuantifica la relación entre dos variables, mientras que la regresión lineal consiste en generar una ecuación (modelo) que, basándose en la relación existente entre ambas variables, permita predecir el valor de una a partir de la otra.

Modelo predictivo: Son un grupo de técnicas que, mediante los campos del aprendizaje automático, la recolección de datos históricos, el Big Data y el reconocimiento de patrones, pretende dar una predicción de resultados futuros, con el objetivo de precisar la toma de decisiones mediante técnicas de análisis de datos.

NoSQL: Son sistemas de gestión de bases de datos y estructuras que permiten almacenar información cuando las bases de datos relacionales dan problemas.

Red neuronal: Con un diseño similar al sistema nervioso y al cerebro humanos, una red neuronal organiza las etapas de aprendizaje para dar a la IA la capacidad de resolver problemas complejos dividiéndolos en niveles de datos. Las redes neuronales aplican la táctica de la división en conjuntos de datos más pequeños para ir superando cada capa de su aprendizaje.

SQL: Se basa en el uso del álgebra y cálculos relacionales para realizar consultas a las bases de datos de manera sencilla. Las consultas se realizan mediante un lenguaje de comandos que permite seleccionar, insertar, actualizar, averiguar la ubicación de los datos, y más.

HR: Es una medida que se relaciona directamente con la frecuencia de las contracciones cardíacas.

Intervalo RR: Es capaz de ofrecer más información que la frecuencia cardíaca, por lo que resultan muy útiles a la hora de evaluar el estado físico de una persona.

LF: son componentes que están alrededor de 0.1 Hz. El poder de producción en la LF depende del tono simpático a causa de la actividad de los barorreceptores.

Sampen: Es una modificación de la entropía aproximada (ApEn), que se utiliza para evaluar la complejidad de las señales fisiológicas de series de tiempo, diagnosticar estados de enfermedad. Sampen tiene dos ventajas sobre ApEn independencia de la longitud de los datos y una implementación relativamente libre de problemas.

pNN50: Es el número de intervalos adyacentes que varían por más de 50 ms expresado en porcentaje.

Capítulo 1 Introducción

El estrés definido por la RAE como la tensión provocada por situaciones agobiantes [1]. El Instituto Nacional de Salud Mental menciona que este afecta a todos y que no siempre es negativo. Destacando el estrés a largo o medio plazo, someterse a grandes niveles de forma continua ocasiona problemas de salud. Mencionan padecimientos cardíacos, presión arterial alta, diabetes y otras enfermedades, incluidos trastornos mentales (depresión o la ansiedad) [2]. Estas constantes alteraciones en el cuerpo comienzan a mermar el sistema inmunitario [3].

La reactividad cardiovascular produce disfunción ventricular izquierda detectada por ecocardiografía. Presenta alteración en el electrocardiograma e incluso elevación en la sangre de marcadores de daño miocárdico. Acciones que surgen por el estrés psicológico cuando sus valores son altos, existe un mayor riesgo de padecer enfermedades cardiovasculares. Los ejes simpaticoadrenal e hipotalámico-hipofisario-adrenal responden ante el estrés agudo con la liberación de grandes cantidades de norepinefrina, epinefrina y cortisol, que le permiten al organismo superar la situación. Aunque el ejercicio físico, actúa como agente protector que evita la secreción de cantidades excesivas de catecolaminas y cortisol cuando el individuo pasa por una situación agobiante. El estrés oxidativo, debido al alto de radicales libres que conlleva, produce a largo plazo una disfunción endotelial responsable de que los vasos sanguíneos no responden ante una situación no placentera. Mediadores liberados actúan durante una respuesta inflamatoria, cuyo estímulo pudiera ser el propio estrés, inducen del mismo modo la disfunción del endotelio, y su intensidad depende de la naturaleza aguda o crónica [4].

En la actualidad, en nuestro país son escasos los instrumentos que permiten monitorear este estímulo de forma cotidiana. Existen pruebas de estrés como el análisis de señales electrocardiográficas (ECG), test de cortisol o por ergometría avalada en área médica. Parte de ellas hacen uso de un dispositivo o personal capacitado, lo que dificulta un monitoreo constante. Aunque existen métodos enfocados en la adquisición como un equipo comercial, brindando un monitoreo sin personal médico. Un ejemplo son las diademas que emplean electroencefalograma (EEG) y frecuencia cardíaca para realizar un registro único. Los relojes inteligentes intentan solucionar el problema mediante el monitoreo continuo de variabilidad cardíaca. Sin embargo, este último sistema utiliza una señal, la fotopletismografía, a partir de la cual se puede hacer una estimación de las variables calculadas con el ECG, tales como la frecuencia y la variabilidad cardíaca. Aunque tiene la característica que, al obtener señales es bajo estimación debido al uso de esta técnica y representa una aproximación de la variable que adquiere [5].

1.1 Estado del arte

Identificar el nivel de estrés es un problema no resuelto, para ello se ha indagado técnicas que nos permitan obtener el mejor resultado. Algunas de estas propuestas buscan detectar los niveles de estrés y ansiedad en pilotos y trabajadores del sector aeronáutico, en los cuales, a partir de un estudio de la respuesta galvánica de la piel, se pudo determinar que la fase de distrés el organismo es incapaz de discernir una solución posible cuando el evento es desfavorable [6]. Debido a que no existe la posibilidad de probar este método en situaciones de la vida real en la que los trabajadores se encuentren con altos niveles de estrés, se ha partido de las emociones que provocan canciones, elegidas de bandas sonoras determinadas. El sensor Grove GSR toma como referencia la sudoración de cada persona en diferentes revisiones efectuadas [7]. Para emplear la mayor cantidad de pruebas, estos datos se han obtenido de tres maneras distintas: la primera es requiriendo quince clips de audio cada uno de 15 segundos. Otra metodología es utilizar siete canciones, pero de mayor duración, 45 s, y la última, una prueba más visual poniendo dos vídeos relacionados con el sector aeronáutico. Tomando lo anterior como referencia, se consideró un protocolo para la adquisición de señales de estrés parecido a la establecido. Con estos datos se han implementado herramientas de aprendizaje automático, obteniendo un modelo de predicción sobre los estados de estrés y ansiedad. Para ello, se realiza un preprocesamiento de los datos, extrayendo las características adecuadas de cada curva. Procedió lo mismo en este proyecto con la diferencia que, lo que se consideró fue la respuesta de un sensor galvánico. La propuesta de este trabajo estudia el comportamiento de los datos brindados por señales EEG y ECG a partir de estadísticos recomendados que se contemplan en la revisión de próximos trabajos [8].

Por diversos factores, no se cuentan con los datos necesarios como para poder crear un modelo predictivo. Una de las razones que cada persona responde de manera diferente a un estímulo. Los datos proporcionados por el sensor son fiables en cierta medida, por último, la cantidad de información presenta demasiadas clases a las que asociar los valores. Finalmente, se ha comprobado cómo cambia la precisión del modelo de predicción cuando en lugar se relacionan 15 variables diferentes de excitación, entre dos posibles estados binarios, 0 o 1. En ambas metodologías, los algoritmos más idóneos para determinar un modelo de predicción, como lo es la técnica de máquinas de vectores de soporte (SVM por sus siglas en inglés) y K vecinos más cercanos (KNN por sus siglas en inglés). Partiendo de esta base, en este trabajo se realizó el análisis de clúster KNN donde verifica si con las señales fisiológicas seleccionadas ocurría el mismo fenómeno al estudiar este tipo de datos.

La tabla 1 resume las características más destacadas con la clasificación de señales de EEG usando técnicas de Inteligencia artificial (IA) y Aprendizaje automático (ML por sus siglas en inglés) encontradas en la bibliografía.

Tabla 1 Estudios previos estrés EEG. Modificado de [8]

ACC: Exactitud	LDA: Análisis Lineal Discriminante
Sen: Sensibilidad	PFC: Corteza prefrontal
NB: Clasificador bayesiano	
KNN: K vecinos más cercano	

Canales	Sujetos	Región cerebral	Clasificador	%	Frecuencia de bandas	Características
4	25	PFC	LDA	Acc: 85.6	4–13 Hz 8–12 Hz 13–30 Hz 25–45 Hz	Frontal Asimetría Alfa, Beta and Gama
22	128	PFC	LDA	Acc: 50	Teta, Alfa, Gama	Intervalo promedio en Gama
22	128	PFC	NB	Acc: 94.6 Sen: 98.3	Delta, teta, Alfa1, Alfa 2, beta1, beta2, beta3, gama1, gama2, gama3	Coherencia y amplitud de asimetría
6	14	Lóbulo prefrontal	SVM KNN LDA	Acc: 89.07 Acc: 72.67 Acc: 70.17	Delta, teta, alfa, beta, gama	El peso del predictor se concentró en la onda alfa
14	1	Lóbulo frontal	RF	Acc: 97.7	Delta (0.5– 3Hz), Teta (3– 8Hz), Alfa (8– 12Hz), Beta (12– 25Hz)	Entropía

En la tabla 1 los estudios que definen el análisis de señal EEG a sujetos expuestos con estímulos de estrés, donde generan técnicas de clasificación con a las frecuencias en las bandas en EEG describiendo la manera más eficiente de usar predictores en su modelo. Recalcando el último estudio donde se destaca enfocar la extracción y análisis particular de cada conjunto es mejor porque se ajusta a sus propios elementos característico. Destacando el número de individuos estudiados, ya que cambian la definición de los factores para poder entrenar el modelo.

Se utilizaron las señales del sensor, dividiendo la base recopilada con los diferentes parámetros a partir de los datos sin procesar. Posterior a obtener subconjuntos de los predictores más precisos para alcanzar una clasificación con alto porcentaje de efectividad. Es necesario tomar referencia del trabajo realizado por Shah NH, Milstein y A, Bagley empleando el programa de selección de características Gain Ratio (semejantes a la que se plantea en este proyecto). Se buscará el algoritmo de clasificación más adecuado mediante los métodos de minería de datos (árboles de decisión, redes bayesianas, k-vecino, etc.) de acuerdo con los patrones identificados del conjunto de la base. Ya que en este proyecto se

estipula que dependiendo del comportamiento de las variables más efectivas que otra. Por lo cual, la diferencia que se desarrollará es que solo se elegirá una técnica que se adecue a lo diseñado al tomar una predicción. Finalmente, proceder a la clasificación, se efectúa una red neuronal convolucional (por sus siglas CNN) siendo este nuestro factor diferencial tendrá una retroalimentación constante, la CNN presentada, el sistema se someterá a nuevos datos, lo que permitirá tener un modelo predictivo y clasificadorio más eficiente [9].

Llegando a este punto es importante resaltar los fundamentos que tomaron de las señales a evaluar son su comportamiento estadístico y extracción de predictores. Razón por la que se consultó el trabajo ejecutado por Steyerberg EW [4] en donde se enfocan en los cambios de los parámetros fisiológicos durante el estrés, evaluando la variabilidad de la frecuencia cardíaca (VFC). Con la diferencia de que no fue un análisis de 6 meses, la investigación que generó un par de horas por un alumno para la adquisición de sus datos y se realizó durante alrededor de una semana. En dicho trabajo se midió la VFC siguiendo su propuesta de rango y las señales que entrenaron pasaron por un preprocesamiento. Se estimó la distancia entre sus características y patrones de cada señal de estos elementos al ser considerado mental o fisiológico. Los parámetros utilizados marcaban el intervalo del nivel de estrés delimitado por la extracción de características. Sirvió al ser una guía para interpretar la señal EEG, añadiendo otro parámetro de correlación al categorizar. Además de corroborar con prueba psicológica y estudio de cortisol, estas variables con el propósito de encontrar una correlación que se plasmara en su modelo predictivo y tendencia del nivel de estrés. Procesar una mayor cantidad de valores al implementar su red neuronal CNN y generar una mejor comprensión en el aprendizaje. Por lo que en el trabajo realizado fue el análisis de una única variable fisiológica usada en su red neuronal planteada obteniendo gran exactitud y efectividad [4].

Como último factor (el análisis de datos cualitativos al añadir al modelo de aprendizaje), presentado por Zainudin Z, et al. implementa un formato semejante de factores (físicos, mentales, de estilo de vida y de personalidad) en alumnos universitarios en diferentes períodos a lo largo de un semestre y modela sus relaciones desempeño-estrés. Se usó como base el algoritmo de aprendizaje automático derivados de los datos recopilados para observar el grado de varianza y distribución. Aunque se utilizaron diferentes tipos de variables, se consideró como base para el modelado de los patrones de aprendizaje necesario. Administrado y guardado en el conjunto de datos correspondiente. Planteó un método de acuerdo mutuo, en lugar de observar la precisión de los resultados, se desarrollaron los parámetros del modelo y los pesos de las características para encontrar tendencias de comportamiento comunes. Este estudio proporcionó información sobre el impacto de diferentes predictores. Se tomó como punto de partida al desarrollar la interfaz y lograr la visualización de valores, nos facilitó su presentación de variables mostrando cómo se resaltaba y ejemplificaba estos [10]. Permite tomar una decisión analítica, construyendo una interpretación del fenómeno de estudio de forma visual (gráficos, esquemas o estadísticos), con el objetivo de captar conceptos difíciles o identificar nuevos patrones. Lo anterior apoyado de un proceso descriptivo, estadístico y técnicas de aprendizaje automatizado. Facilitó la comprensión y aplicación de los resultados de cualquier estudio de investigación, en nuestro caso la identificación de la tendencia y categorización del nivel de estrés. El apartado de análisis estadístico es básico en el protocolo de investigación, por lo que siempre debe estar presente.

Capítulo 2 Objetivos

2.1 Objetivo general

Implementar un sistema para la categorización del nivel de estrés a partir de la clasificación, almacenamiento, análisis e interpretación de EEG y ECG, mediante la utilización de modelos estadísticos, algoritmos de aprendizaje automático y técnicas de minería de datos.

2.2 Objetivo específico

- Implementar un dispositivo de monitoreo en tiempo real con sensores no invasivos que permitan la adquisición de señales ECG y EEG, una tarjeta de desarrollo de tipo Raspberry Pi 4 Modelo B y como tarjeta de desarrollo auxiliar un Arduino Uno. Que cuente con una interfaz gráfica que facilite uso al público en general.
- Desarrollar un procesamiento digital con filtros y técnicas de reducción de ruido que permitan garantizar el acondicionamiento óptimo del conjunto de señales analógicas propuestas ECG y EEG. Y además de garantizar un acoplamiento de impedancias.
- Diseñar e implementar una red neuronal capaz de clasificar el estrés mental multinivel con una señal EEG y delimite potencia del ritmo alfa en la corteza prefrontal como un índice adecuado para su posterior clasificación.
- Implementar algoritmo que detecte el índice de variabilidad cardiaca a través de la identificación de las características de la señal ECG que permite un modelado multinivel que delimite la variabilidad de la tasa para la identificación del estado de estrés.
- Diseñar una base de datos SQL para tener un óptimo manejo de consulta de los resultados obtenidos. Estableciendo una conexión con Python para llenar cada tabla con sus respectivos datos resultantes de la categorización.
- Realizar un muestreo para adquirir las señales ECG y EEG de la muestra estadística de los estudiantes. Delimitando a veinte personas como mínimo de datos requeridos de manera que se genere una base de datos con mil datos inferenciales como mínimo que permitan la identificación de patrones de estímulo del nivel de estrés. Aumentando la muestra estadística a 200 usando la base de datos del MIT.
- Elaborar un informe que muestre de manera global un análisis e interpretación de datos apoyado de un modelo estadístico congruente a los datos obtenidos en la etapa de muestreo. Mostrar qué es relevante en los datos y posteriormente hacer un uso correcto de esa información para evaluar resultados probables y generar de esa manera un modelo predictivo con técnicas de aprendizaje automático.

Capítulo 3 Planteamiento del problema

De acuerdo con los trabajos de investigación, la mayor parte de ellos se concentraron en la detección de estrés actual y solo algunos consideran importante el seguimiento en un intervalo de tiempo mayor. Colocando en segundo plano tener un historial que indique el número de veces que una persona experimenta estrés ante cualquier estímulo. Surgiendo como desventaja de una pérdida de información que pudiera ser presentada a un profesional del área médica o psicología y ellos decidir si tiene valor como indicador o herramienta que aporte a la revisión retroalimentada.

Un beneficio que se pierde con esa información es, que no se puede diseñar un modelo predictivo que describa la tendencia que tiene un individuo al estar estresado. Posterior a obtener datos relevantes que puedan ayudar a interpretar esa tendencia, un problema visto desde el área de análisis probabilístico. Entre las limitaciones del proyecto, la muestra no es suficiente para validar el algoritmo de categorización y eso puede ocasionar un problema de clasificación lineal, porque al realizar el modelo se requiere de muchas variables para poder desarrollarse y presentar una mayor eficiencia, por lo que se buscará en bases de datos de instituciones académicas. Además, el modelo depende de una alta calidad en los datos, debido a que esto puede ser una limitante en el entrenamiento; por lo que se buscaría solucionar estas problemáticas al momento de implementar y poder identificar los patrones adecuados en las señales ECG y EEG que se analizaron.

El estrés se puede interpretar desde tres perspectivas: la ambiental, que se centra en los eventos vitales; la psicológica, que implica la experiencia subjetiva y respuesta emocional frente a los estresores, estímulos agobiantes; y la biomédica, que estudia el accionar y los sistemas fisiológicos involucrados en el afrontar los asuntos vitales. Buscando desarrollar una clasificación analizando los tres factores para conocer a través de datos cuantificables el nivel de estrés que tiene el sujeto de una manera que recurra al aprendizaje no supervisado y al análisis de las variables. Al delimitar predictores de nuestro modelo, finalizar con la categorización en agrupación de las señales con su respectiva interpretación.

Capítulo 4 Justificación

La enfermedad del siglo, para las organizaciones estadounidenses, el costo anual de tratar el estrés laboral asciende a 220 mil millones de dólares en el tratamiento de enfermedades físicas, emocionales y comportamientos que afectan su desempeño. En marzo de 2020 se iniciaron encuestas e investigaciones con el propósito de intuir el impacto psicológico del confinamiento social en la salud mental del cuerpo estudiantil de diversas universidades mexicanas. La encuesta realizada por González Jaimes, notan rasgos depresivos 4.9 %, psicosomáticos 5.9 %, a sociabilidad 9.5 %, estrés 31.92 % y cansancio mental 36.5 % en una muestra de 644 estudiantes universitarios mexicanos. Con los datos antes mencionados, surge la finalidad de apoyar en el reconocimiento temprano del estrés que ocurre de forma recurrente. En la actualidad, existen soluciones para identificarlo, como el uso de un electroencefalograma (EEG) evaluar el estrés de los individuos mediante el análisis de las ondas cerebrales en los dominios clínicos [11].

Existen dispositivos EEG y ECG portátiles con diferentes parámetros que ayudan en su identificación. Sin embargo, lo diferencial del trabajo es reconocer estrés al que una persona se ve sometida en su rutina. Recurriendo a señales EEG y ECG como principales señales para entrenar nuestro modelo para la clasificación. En cuanto a las implicaciones prácticas, el trabajo no resolverá una problemática de manera directa, será un indicador de frecuencia y duración a un estímulo que provoque estrés. Analizando la tendencia y comportamiento de sus datos para generar un modelo de aprendizaje automático. Concluyendo en una categorización del estrés (tomando como factores a delimitar este punto; frecuencia, intervalo de tiempo y comportamiento de las señales biomédicas).

El valor teórico, la finalidad es desarrollar métodos para identificar a las personas en riesgo de sufrir del estrés y servir como medida de prevención. Busca respaldar intervenciones tempranas con el monitoreo y como reto implementar tomando las variables fisiológicas para generar un modelo matemático que describa su comportamiento. Existen modelos con otros tipos de datos, por ejemplo, el aprendizaje de un algoritmo predictivo, donde se evaluó el efecto de la pandemia sobre los niveles de estrés en una muestra, 2053 adultos italianos, y se caracterizó a los individuos más vulnerables de la base de características socio demográficas y rasgos psicológicos estables. Brindando un valor agregado a nuestra herramienta al ser implementado, permite un almacenamiento, captura y análisis de datos de manera particular de cada persona buscando que no exista un sesgo en el modelo como ocurrió en este trabajo. Buscará proporcionar conocimientos y técnicas estadísticas con la finalidad de que pueda convertirse en una herramienta de detección útil [11]. Existiendo proyectos con esta misma propuesta tener un modelo de aprendizaje profundo (DL por sus siglas en inglés), enfocado en el aprendizaje automático con la capacidad de detectar, aunque solo de fase binaria [12]. En este trabajo se plantea una resolución multi nivel a la clasificación de estrés destacando este como nuestro diferenciador en el proyecto realizado. A razón de lo anterior, el objetivo es categorizar el nivel de estrés de una forma no subjetiva.

Capítulo 5 Marco Teórico

5.1 Señal EEG

Las bandas de las señales EEG son nombradas con cierta nomenclatura; según su oscilación las que caen debajo de 1 Hz y por encima de 20 Hz son consideradas como artefactos. Hay seis tipos de señales que son Delta (δ), Teta (θ), Alpha (α), Mu (μ), Beta (β) y Gamma (γ). Delta (δ) tiene una amplitud variable y cae en el rango de 1-4 Hz. Delta parece ser la más alta en amplitud y la onda más lenta asociada con los estados del sueño profundo y despertar. Teta (θ) tiene una amplitud que es mayor que 20 μ V y está dentro del rango de 4-7 Hz. Teta se genera con la inspiración inactiva, creativa, inconsciente material, somnolencia y meditación profunda. Alpha (α) es una onda con amplitud de 30-50 μ V y velocidad de cambio entre 8 y 13 Hz, es usualmente asociada con la relajación, concentración y algunas veces atención. Mu (μ), se encuentra en el rango de frecuencia de ondas alfa donde la amplitud registrada sobre la corteza motora es máxima, ordinariamente, es asociada con la supresión, la cual indica que las neuronas motoras están funcionando. Beta (β) se asocia con alerta, pensamiento y concentración activa y cae en el rango entre 12 y 30 Hz. Finalmente, Gamma (γ) podría ser detectada en la corteza somatosensorial con frecuencia mayor a 30 Hz. También se muestra durante la coincidencia de memoria a corto plazo de objetos reconocidos, sonidos o sensaciones táctiles.

Un electroencefalograma es una prueba que detecta la actividad eléctrica del cerebro mediante pequeños discos metálicos (electrodos) fijados sobre el cuero cabelludo. Las neuronas cerebrales se comunican a través de impulsos eléctricos y están activas todo el tiempo, incluso mientras duermes. Esta actividad se manifiesta como líneas onduladas en un registro de electroencefalograma. Es uno de los estudios principales para diagnosticar la epilepsia. Un electroencefalograma también puede cumplir una función en el diagnóstico de otros trastornos cerebrales [7].

El sistema de posicionamiento 10-20 es un estándar para la grabación de señales EEG, la numeración 10-20 se refiere a puntos de referencia anatómicos específicos o distancia entre electrodos; representando el 10-20 % de los perímetros de la cabeza de adelante hacia atrás o de derecha a izquierda. Otra alternativa, son los sistemas 10-10 y 10-5 que según Kumar y Bhuvaneswai son un estándar válido [7]. Durante la toma de muestras el EEG es digitalizado normalmente con un convertidor analógico a digital (ADC) de 12 bits con frecuencia de muestreo que va desde 100 Hz a varios cientos de Hercios.

5.1.1 Potencial en neuronas

El sistema nervioso está compuesto por millones de neuronas, que están conectadas entre sí para formar una red compleja. Las células interactúan a través de impulsos eléctricos a través de dendritas y axones. Las dendritas son procesos de ramificación alargados que se extienden desde el citoplasma del cuerpo celular. Estos proporcionan un área de recepción para transmitir pulsos electroquímicos graduales al cuerpo de la celda. El axón es una extensión larga que puede conducir impulsos llamados potenciales de acción desde el cuerpo celular hacia la periferia. El axón termina cerca de otra célula o, en algunos casos, en algún lugar existe un punto de contacto. Una vez que el potencial de acción llega al final del axón,

estiman o inhiben directa o indirectamente a otra célula. El potencial eléctrico a través de la membrana puede ser medido penetrando la célula con un electrodo. Una neurona en reposo tiene una diferencia de voltaje de aproximadamente -60 mV a través de la membrana y este valor es su potencial de reposo. Los potenciales de acción casi siempre se detienen al final del axón, donde estimulan la liberación de neurotransmisores que afectan a la siguiente célula; esta conexión se denomina sinapsis. Como se muestra en la figura 1 donde referencia la gráfica de potencial de acción que en este se genera [7].

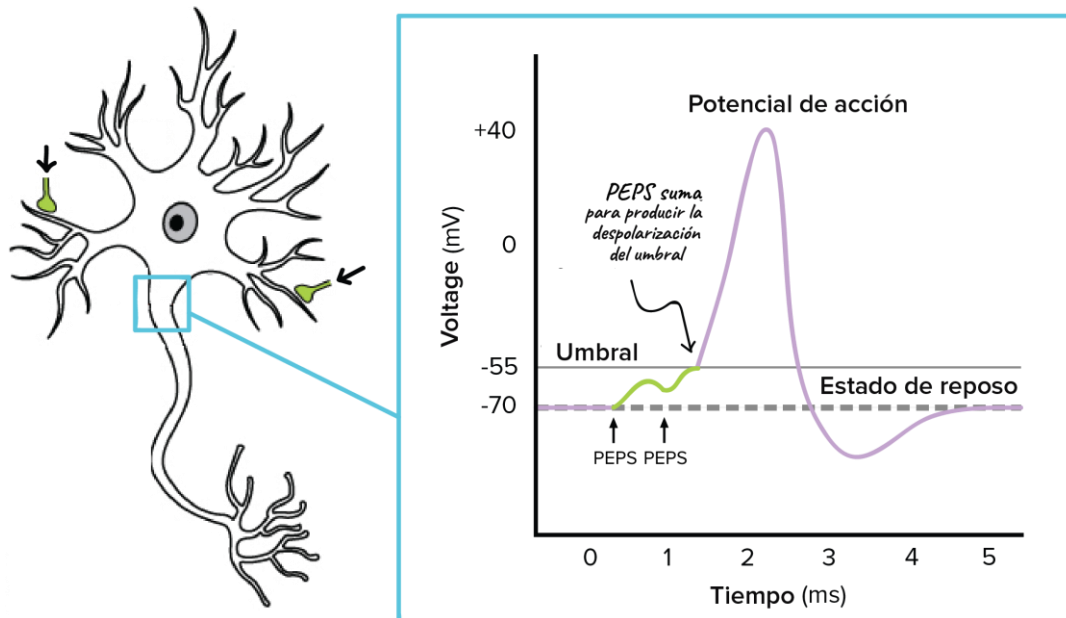


Figura 1 Sinapsis. Recuperado [28]

Recordando que el sistema nervioso central (SNC) está formado por el encéfalo y la médula espinal. El SNC recibe aferencias provenientes de neuronas sensoriales y dirige la actividad de neuronas motoras que inervan músculos y glándulas. El cerebro está compuesto por dos hemisferios, el hemisferio derecho y el izquierdo, el interior consta de una gran área llamada cuerpo duro. El cuerpo calloso es el principal haz de axones. Conecta funcionalmente los dos hemisferios cerebrales. Los hemisferios cerebrales muestran especialización de ciertas funciones y fenómenos a esto se le llama lateralización del cerebro.

Cada hemisferio está separado por profundas ranuras o grietas, formando cinco hojas, cuatro de las cuales son visibles desde la superficie. Son los lóbulos frontal, parietal, temporal y occipital. Los lóbulos insulares se localizan en el plano profundo y están cubiertos por lóbulos parciales. Los lóbulos frontal, parietal y temporal. El lóbulo frontal es la parte frontal de cada hemisferio cerebral. Una fisura profunda llamada surco central separa el lóbulo frontal del lóbulo parietal. La circunvolución anterior central participa en el control motor y se encuentra en el lóbulo frontal antes del surco central. Debido a su papel en la regulación muscular, los cuerpos celulares de las interneuronas que se encuentran allí se denominan neuronas motoras superiores. La circunvolución posterior central se localiza directamente detrás del surco central del lóbulo parietal. La percepción somatosensorial depende de los receptores de la piel, los músculos, los tendones y las articulaciones.

Corteza Cerebral

El papel funcional de las diferentes áreas de la corteza cerebral es descrito de la siguiente manera: las áreas principales, que incluyen la corteza motora primaria y las áreas que reciben estímulos sensoriales: corteza somatosensorial primaria (lóbulo parietal) para información sensorial, corteza auditiva primaria (lóbulo temporal) y para la secundaria lóbulo visual (lóbulo occipital). Las áreas secundarias, que corresponden a información elaborada del procesamiento que puede ser plurimodal y áreas asociativas, cuyas funciones son funciones cognitivas y atencionales, las cuales prestan atención a los estímulos para identificarlos. Las funciones cognitivas son procesadas en tales áreas.

5.1.2 Sistema de conducción EEG

Tratar de comprender un sistema complejo como el cerebro requiere estudiar su evolución, anatomía y estructura funcional, así como el entorno en el que trabaja, por lo que es necesario analizar cómo interactúan estos múltiples factores o variables para producir patrones dinámicos. La conectividad muestra estos patrones de conexión se pueden representar mediante la interacción de dos componentes simples: nodos y enlaces. A través de esta abstracción, se puede generar una red conectada, y luego se puede estudiar el comportamiento de los elementos de la red en función de cambios dinámicos en el cerebro, grupo global o local. Se han descubierto parámetros que caracterizan redes complejas de varios patrones de conexión cerebral. Al aplicar las características de las redes del mundo pequeño para modelar redes interconectadas a gran escala, redes locales y globales, se puede estudiar el comportamiento funcional del cerebro humano.

Su aplicación comenzó a producir resultados importantes en el conocimiento de las funciones normales, por ejemplo, en el estudio del desarrollo cerebral, circuitos que interfieren con el sueño, el movimiento de manos o pies, o la imaginación al respecto, y por lesión cerebral, epilepsia, enfermedad de Parkinson, enfermedad de Alzheimer y esquizofrenia. Un gráfico es una representación de una red de objetos, derivados de nodos (vértices) vinculados entre sí a través de conexiones (bordes). Según el valor estimado de la fuerza de la conexión, la conexión puede ser de tipo binario (conectada, no conectada) o reconectada. El EEG se puede analizar a partir del gráfico, que utiliza cada electrodo del componente como nodo, el borde que conecta todos los electrodos entre sí (gráfico completo) y el peso derivado de la medida de estimación de conectividad, tratando de cuantificar cualquier señal de relación entre los electrodos [11].

5.2 Electrocardiograma y señal ECG

El electrocardiograma es un registro gráfico de los potenciales eléctricos generados en el corazón durante ciclos cardíacos [12]. Un electrocardiograma es de gran importancia en el área cardiovascular, ya que entrega información muy útil acerca del funcionamiento del corazón. Un electrocardiograma es una representación gráfica de la señal ECG, la que consiste en una línea base y varias deflexiones y formas de ondas variables en el tiempo. Ubicando electrodos en diversas posiciones del cuerpo humano y conectando dichos electrodos a un aparato electrocardiográfico, es posible realizar mediciones de las variaciones del potencial eléctrico cardíaco, cuya visualización es conocida como electrocardiograma. Una disposición específica de los electrodos en el cuerpo humano recibe el nombre de

derivación [13]. Las derivaciones más importantes reciben el nombre de derivaciones bipolares, derivaciones aumentadas y derivaciones precordiales unipolares.

5.2.1 Ancho de banda características ECG

La amplitud de la señal a recolectar está entre ± 5 V. Esto se debe al uso de canales de instrumentos analógicos de ECG con estas características. El ancho de banda de la señal será tan alto como 150 Hz y puede estar limitado a 17 Hz, 40 Hz o 100 Hz. La frecuencia de corte baja siempre se establece en 0,05 Hz. La American Heart Association (AHA, de la Anglo-American Heart Association) recomienda 0,05 Hz, principalmente debido a la distorsión de la información relacionada con el segmento ST, que se produce en una frecuencia de corte superior a 0,1 Hz. Todo esto recordando que la señal ECG tiene un espectro frecuencial que va desde los 0 Hz a los 100 Hz. Este se representa en la figura 2 donde se observa el comportamiento y tendencia de la amplitud y frecuencia en la banda de ECG [13].

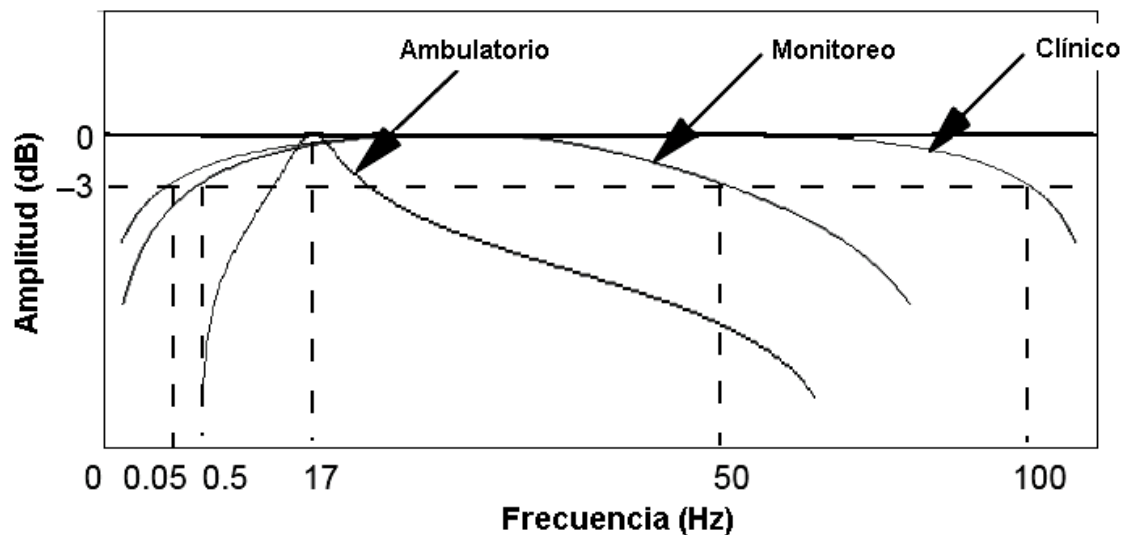


Figura 2 Representación de ancho de banda ECG. Recuperado [29]

En cuanto a la frecuencia de muestreo tiene que ser superior a la utilizada usualmente en sistemas ambulatorios de registro de señales biomédicas, más conocidos como Holter, que es de 128 Hz. La conversión analógica-digital tendrá entonces dos valores posibles: 500 muestras o 1000 muestras, ambas con una resolución mínima de 12 bits. Con ello se garantiza el cumplimiento del Teorema de Nyquist, aún para la frecuencia de corte a las altas frecuencias, de 150 Hz. Sin embargo, trae consigo un aumento en el tamaño del archivo de registro y que estos contengan un mayor nivel de ruido.

5.2.2 Sistema de conducción cardíaco

El sistema de conducción cardíaco son las estructuras desde donde se produce y se trasmite el estímulo eléctrico que permite la contracción del corazón. Sus principales elementos son el nodo sinusal, el nodo auriculoventricular (nodo AV), el haz de His y las fibras de Purkinje. En un latido normal, el impulso eléctrico es generado por el nodo sinusal, desde donde se propaga a ambas aurículas, provocando la contracción auricular. Mediante las vías preferenciales auriculares el impulso llega al nodo AV que, tras retrasar el impulso, lo trasmite al haz de His y este lo propaga a todo el miocardio por las fibras de Purkinje. Lo anterior se representa en la figura 3.

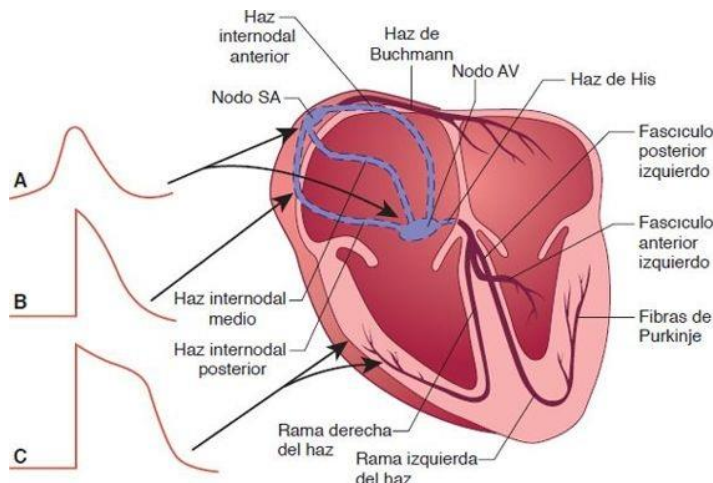


Figura 3 Sistema de conducción cardíaco. Recuperado [30]

5.3 La Variabilidad de la Frecuencia Cardíaca (VFC)

Es un parámetro indicador del nivel de salud cardiovascular, de manera que su análisis permite un estudio no invasivo de la actividad del sistema nervioso autónomo (por sus siglas SNA) sobre el nodo sinusal: una alta VFC es señal de un buen funcionamiento y adaptabilidad del SNA, considerado como un estado saludable mientras una baja VFC significa un funcionamiento anormal o una adaptabilidad insuficiente del SNA, lo que es indicador de mala salud. La VFC se define como la variación que ocurre en el intervalo de tiempo, entre dos latidos cardíacos consecutivos. Se obtiene identificando el momento donde aparece cada onda R y se calcula el tiempo transcurrido entre dos ondas R consecutivas. [14]

5.4 Estimación de nivel de estrés

Estimación del estrés ha sido la evaluación continua y en tiempo real del grado de ansiedad o activación experimentado por los participantes sin la necesidad de recurrir a medidas de autoinforme. Esta aportación es de gran utilidad en estudios o contextos en los que la evaluación explícita resulta difícil o imposible, ya sea por las características de la tarea o por circunstancias específicas de la población a evaluar. Este es el caso, por ejemplo, de un estudio llevado a cabo con pilotos que eran sometidos a pruebas de simulación de vuelos. Su estado de activación y ansiedad fue satisfactoriamente evaluado mediante el uso de esta metodología durante toda la tarea. De modo semejante, el uso de este tipo de indicadores

puede ser muy adecuado cuando se trata de registrar la actividad mental (p. ej., interés o esfuerzo) en participantes que, como los bebés, no se comunican verbalmente [14].

5.4.1 Sistemas de instrumentación para filtrado de señales analógicas

El método de filtrado que presenta el mejor rendimiento frente al ruido es Chebyshev. Pero tienen el problema de introducir distorsión en la señal debido al sobre impulso. En otras palabras, existe una compensación entre la eliminación de ruido y la distorsión, esta última limita la frecuencia de las señales de pulso que se pueden filtrar debido al tiempo de estabilización. Una mayor eliminación de ruido significa una mayor distorsión; menos distorsión significa menor inmunidad al ruido. En vista de esto, se recomienda utilizar la tecnología de eliminación de sobretensiones proporcionada por la ingeniería de control, combinada con el filtro con el mejor desempeño anti-ruido, y buscar un nuevo método de diseño de filtro analógico para ajustar la señal de pulso que no cambiará o lo menos posible. Cambie la forma de la señal de pulso. Tecnología para eliminar sobretensiones. En un entorno industrial, el filtrado es fundamental para eliminar el ruido y la interferencia combinados con la señal del sensor. Para realizar esta función, existen filtros analógicos, que pueden ser activos o pasivos. Tome la instrumentación industrial como ejemplo para el caso de la instrumentación industrial, los filtros activos son los más utilizados, debido a que se puede prescindir de los inductores cuando se requieran implementaciones de alto orden

5.5 Redes neuronales

La capacidad de aprendizaje adaptativo es una de las características más atractivas de redes neuronales. Esto es, aprenden a llevar a cabo ciertas tareas mediante un entrenamiento con ejemplos ilustrativos. Como las redes neuronales pueden aprender a diferenciar patrones mediante ejemplos y entrenamientos, no es necesario elaborar modelos a priori sin necesidad de especificar funciones de distribución de probabilidad. Las redes neuronales son sistemas dinámicos auto adaptativos. Son adaptables debido a la capacidad de autoajuste de los elementos procesales (neuronas) que componen el sistema. Son dinámicos, pues son capaces de estar constantemente cambiando para adaptarse a las nuevas condiciones. En el proceso de aprendizaje, los enlaces ponderados de las neuronas se ajustan de manera que se obtengan ciertos resultados específicos. Una red neuronal no necesita un algoritmo para resolver un problema, ya que ella puede generar su propia distribución de pesos en los enlaces mediante el aprendizaje. También existen redes que continúan aprendiendo a lo largo de su vida, después de completado su período de entrenamiento. La función del diseñador es únicamente la obtención de la arquitectura apropiada. No es problema del diseñador el cómo la red aprenderá a discriminar. Sin embargo, sí es necesario que desarrolle un buen algoritmo de aprendizaje que le proporcione a la red la capacidad de discriminar, mediante un entrenamiento con patrones. [16]

5.6 Modelos de regresión multivariable

Los modelos de regresión multivariables se utilizan ampliamente en la investigación de ciencias de la salud. Con frecuencia, el objetivo en la recolección de datos obedece al afán de explicar las interrelaciones que existen entre ciertas variables o a determinar los factores que afectan a la presencia o ausencia de un episodio adverso determinado. Es ahí donde los

modelos de regresión multivariantes pasan a ser un instrumento útil, al suministrar una explicación matemática simplificada de dicha relación. El objetivo final será obtener un modelo simplificado que tenga sentido desde una perspectiva biológica, se atenga estrechamente a los datos disponibles y aporte predicciones válidas al aplicarlo a datos independientes. Actualmente, investigadores con una formación no muy extensa en estadística a menudo se dedican al análisis de sus datos, en gran parte debido a la disponibilidad y la facilidad de uso de los paquetes informáticos más comunes [17].

5.7 Árbol de decisión

Son modelos predictivos formados por reglas binarias (sí/no) con las que se consigue repartir las observaciones en función de sus atributos y predecir así el valor de la variable respuesta. Muchos métodos predictivos generan modelos globales en los que una única ecuación se aplica a todo el espacio muestral. Cuando el caso de uso implica múltiples predictores, que interactúan entre ellos de forma compleja y no lineal, es muy difícil encontrar un único modelo global que sea capaz de reflejar la relación entre las variables. Los métodos estadísticos y de machine learning basados en árboles engloban a un conjunto de técnicas supervisadas no paramétricas que consiguen segmentar el espacio de los predictores en regiones simples, dentro de las cuales es más sencillo manejar las interacciones. Es esta característica la que les proporciona gran parte de su potencial [18].

5.8 Clúster

El término clustering hace referencia a un amplio abanico de técnicas cuya finalidad es encontrar patrones o grupos (clusters) dentro de un conjunto de observaciones. Las particiones se establecen de forma que, las observaciones que están dentro de un mismo grupo son similares entre ellas y distintas a las observaciones de otros grupos. Se trata de un método de aprendizaje no supervisado, ya que el proceso no tiene en cuenta a qué grupo pertenece realmente cada observación (si es que existe tal información). Esta característica es la que diferencia al clustering de los métodos de clasificación en el que sí emplea la verdadera clasificación durante su entrenamiento [17].

5.9 Análisis por factorial

El análisis factorial (FA) es un método de análisis de datos exploratorio que se utiliza para buscar factores subyacentes influyentes o variables latentes de un conjunto de variables observadas. Ayuda en la interpretación de datos al reducir el número de variables. Extrae la varianza máxima común de todas las variables y las coloca en una puntuación común.

El análisis factorial se utiliza ampliamente en investigación de mercado, publicidad, psicología, finanzas e investigación operativa. Los investigadores de mercado utilizan el análisis de factores para identificar a los clientes sensibles al precio, identificar las características de la marca que influyen en la elección del consumidor y ayudar a comprender los criterios de selección de canales para el canal de distribución [17].

5.10 Red neuronal Autoencoder

Un autoencoder es una red neuronal utilizada para aprender codificaciones (p.e. reducción de dimensionalidad, compresión, cifrado, etc) eficientes. Son redes neuronales con el objetivo de generar nuevos datos primero comprimiendo la entrada en un espacio de variables latentes y luego reconstruyendo la salida con base en la información adquirida [17].

5.11 La Escala de Estrés Percibido (EEP)

Es una de las escalas más conocidas para la medición de la respuesta psicológica general frente a los estresores. La evaluación psicométrica de EEP en diferentes ámbitos y poblaciones permite que se cuente en la actualidad con varias versiones de esta herramienta de medición. Entre las diferentes versiones se encuentra la original, compuesta por 14 incisos (EEP-14), así como otras que son el resultado de un proceso continuo de refinamiento; en este último grupo, en el contexto mundial, las dos más conocidas son una versión de 10 (EEP-10) y otra de 4 reactivos (EEP-4), con traducciones en diferentes idiomas, incluido el español [35].

Capítulo 6 Desarrollo experimental

En el Anexo A añade la carta de participación de retroalimentación de este trabajo desde el punto de vista del área psicología para su aplicación. La meta buscada fue implementar una investigación práctica para comprobar si el proceso de clasificación tiene congruencia a efectos de su cateterización multi nivel.

6.1 Sujetos de estudio y tamaño de la muestra

El protocolo de consentimiento a cada persona que fue partícipe en el protocolo de datos se proporcionó en que consistió y firmaron una carta de consentimiento agregada en el Anexo A. Buscando como objetivo en todo momento, él respetó al Tratado de Helsinki los principios éticos para las investigaciones médicas en seres humanos. Para llevar a cabo el protocolo de estimulación para generar un estímulo de estrés, se siguió el protocolo tarea de estrés por imágenes de (MIST, por sus siglas en inglés). Se define por ser un protocolo derivado del segmento de desafío mental de la prueba de estrés social de Trier, diseñado para inducir estrés psicológico moderado mencionado por Dedovic. Este procedimiento consiste en una serie de desafíos aritméticos mentales computarizados (con limitación de tiempo y manipulados para estar al límite de la capacidad mental del individuo, complementados con la presentación del promedio de desempeño del individuo y el desempeño esperado), en combinación con componentes de evaluación social.

6.1.1 Protocolo MIST

Este protocolo ha demostrado incrementar los niveles de cortisol en saliva de manera significativa, y ha sido útil en los estudios que requieren el uso de técnicas como la resonancia magnética funcional (por sus siglas en inglés fMRI) y la tomografía por emisión de positrones (por sus siglas en inglés PET). En este caso, el modelo demuestra ser efectivo en la activación del eje hipotalámico-hipofisario-adrenal, sin embargo, el componente de evaluación social puede generar resultados diferenciales, dependiendo de las características del participante, por su naturaleza subjetiva (esto se debe al implementar la escala de medición del estrés percibido).

Una ventaja importante de este protocolo es su adaptabilidad a los equipos que requieren que escáner como fMRI y PET. El resultado de la implementación del protocolo MIST resultó en la recopilación de la siguiente base de datos compuesta por bases públicas y de elaboración propia. Señalando un muestreo de veinte personas y una muestra estadística general con las bases de datos del MIT y Kaggle de 110 personas.

El protocolo MIST consiste en una sesión de entrenamiento, realizada fuera de la unidad de imágenes, y una sesión de prueba, durante la cual se adquirió imágenes funcionales. Dependiendo del método de imagen utilizado, se implementó hasta 3 condiciones diferentes (reposo, control y experimental). También se comprobó que con este protocolo el uso de imágenes funcionales para investigar los efectos de percibir y procesar el estrés psicosocial en el cerebro humano aumenta el nivel de cortisol en las personas.

6.1.2 Muestreo y sujetos de estudio

En este cúmulo de sujetos se generalizaron sus datos, es decir, no se sabe la edad de la persona, estado de salud o la condición social que la persona tenía al momento de realizar el protocolo. Al proceder el muestreo de las 20 personas tampoco se consideraron estos factores para que no se viera afectados la base de datos pública recopilados. La base de datos elaborada de las siguientes señales EEG y ECG.

Tabla 2 Muestras [24][25][26] [27] de proyectos de investigación y muestras adquiridas por los dispositivos y datos de prueba adquisición propia

<i>Tipo de señal</i>	<i>Entrenamiento</i>	<i>Prueba</i>
EEG	369,290	369,290
ECG	4,537	1,162

- **Total: 6,154 horas en adquisición de señales ECG**
- **Total: 125 horas en adquisición de señales EEG**

Los datos mostrados en la Tabla 2 son el resultado de generar una recopilación de bases de datos públicas combinada con señal adquirida que se acoplaron a las generalidades de estas. Estas no especifican edad, género, antecedentes médicos u otro dato relevante en la elaboración, por lo que esos factores no fueron considerados y tampoco sesga la información obtenida.

Posterior a adquirir las señales se implementó la Escala de Estrés Percibido (EEP)

Tabla 3 Escala de Estrés Percibido (EEP) [36]

Durante el último mes	Nunca	Casi nunca	De vez en cuando	A menudo	Muy a menudo
1. ¿Con qué frecuencia ha estado afectado por algo que ha ocurrido inesperadamente?	0	1	2	3	4
2. ¿Con qué frecuencia se ha sentido incapaz de controlar las cosas importantes en su vida?	0	1	2	3	4
3. ¿Con qué frecuencia se ha sentido nervioso o estresado?	0	1	2	3	4
4. ¿Con qué frecuencia ha estado seguro sobre su capacidad para manejar sus problemas personales?	4	3	2	1	0
5. ¿Con qué frecuencia ha sentido que las cosas le van bien?	4	3	2	1	0
6. ¿Con qué frecuencia ha sentido que no podía afrontar todas las cosas que tenía que hacer?	0	1	2	3	4
7. ¿Con qué frecuencia ha podido controlar las dificultades de su vida?	4	3	2	1	0
8. ¿Con qué frecuencia se ha sentido que tenía todo bajo control?	4	3	2	1	0
9. ¿Con qué frecuencia ha estado enfadado porque las cosas que le han ocurrido estaban fuera de su control?	0	1	2	3	4
10. ¿Con qué frecuencia ha sentido que las dificultades se acumulan tanto que no puede superarlas?	0	1	2	3	4

Se guió este protocolo a partir de la teoría transaccional del estrés (Lazarus y Folkman, 1984), Cohen, Kamarck y Mermelstein (1983) construyeron la “Escala de estrés percibido” (EEP)

(Perceived Stress Scale, PSS), instrumento que estima el grado en que las situaciones de la vida son valoradas como estresantes por las personas. Posteriormente, se almacenó en la base de datos añadiendo otro factor de predicción al modelo sirvió para cuantificar por dos factores a partir de estrés percibido por sujeto y por las señales fisiológicas EEG y ECG.

6.2 Instrumentos

El dispositivo se usó para adquirir las señales EEG (Mindflex) y para adquirir la señal ECG (Sensor Ad8232) y corroboró con fotopletimografía (en caso de que requiriera una medición práctica sin uso de electrodos). Como tarjeta de desarrollo se emplea una Raspberry Pi 4 y como apoyo para interpretación de datos a través de librerías utilizó la tarjeta de desarrollo Arduino Uno.

6.2.1 Mindflex

Microcontrolador que usa: (La hoja de especificaciones se añade en el Anexo C)

El TGAM es el módulo ASIC del sensor de ondas cerebrales principal de NeuroSky diseñado para aplicaciones de mercado masivo. El TGAM procesa y produce la frecuencia de EEG espectros, calidad de la señal de EEG, EEG sin procesar y tres medidores NeuroSky eSense atención; meditación; y cejas. Con electrodos secos simples, este módulo es excelente para usar en juguetes, videojuegos y dispositivos de bienestar debido a su bajo costo en comparación a otros dispositivos en el mercado.

Características:

- TGAM: ThinkGear ASIC Module
- Se conecta directamente al electrodo seco (a diferencia de los sensores húmedos médicos convencionales)
- Un canal EEG con tres contactos: EEG; ÁRBITRO y GND
- Ajuste inadecuado detectado a través de la advertencia de "Calidad de señal deficiente" del ASIC para reiniciar si no está en la cabeza durante cuatro consecutivos segundos, o si recibe una mala señal durante siete segundos consecutivos
- Tecnología de filtrado avanzada con alta inmunidad al ruido
- Bajo consumo de energía adecuado para aplicaciones portátiles a batería
- Consumo máximo de energía 15 mA a 3,3 V
- Salida de datos de EEG sin procesar a 512 bits por segundo

Datos brindados:

- Procesamiento y salida de espectros de potencia EEG (Alfa, Beta, etc.)
- Procesamiento y salida del medidor eSense patentado por NeuroSky para atención, meditación y otros medidores futuros
- Análisis de la calidad de la señal de EEG / ECG (se puede utilizar para detectar un contacto deficiente y si el dispositivo está fuera de la cabeza)

Especificaciones

- Frecuencia de muestreo de 512 bits por segundo
- Rango de frecuencia de 3-100Hz
- Protección ESD: Descarga de contacto de 4kV; Descarga de aire de 8kV
- Consumo máximo de energía: 15 mA a 3,3 V
- Voltaje de funcionamiento 2,97 ~ 3,63 V

Interfaz de salida estándar UART (serial):

- Tasa de baudios de salida de 1200, 9600, 57600
- 8 bits
- Paridad: ninguna
- Bit de parada: 1

Opciones configurables Filtro de ruido de CA

- 50 Hz y 60 Hz

Salida y tasa de baudios:

- 1200 baudios con atención, meditación, potencias de EEG (alfa, beta, etc.) y mala calidad de señal
 - 9600 baudios con atención, meditación, poderes de EEG (alfa, beta, etc.) y mala calidad de señal
 - 57600 baudios con atención, meditación, poderes de EEG (alfa, beta, etc.) y señal deficiente
- Onda de EEG de calidad y sin procesar

Electrodos especificaciones que puede usar:

- Área de superficie máxima de ~ 150 mm² (pero menos área de superficie es óptima)
- Ag / AgCl, acero inoxidable, oro y / o plata (tanto el material sólido como el plateado funcionan)
- Electrodo de EEG ubicado sobre el ojo izquierdo o derecho en la frente
- Electrodos de tierra y de referencia ubicados detrás de la oreja o en el lóbulo de la oreja
- Tener suficiente presión para evitar el movimiento, con un mínimo de 0,8 PSI

Cable de los electrodos:

- Longitud de menos de 12 pulgadas, cuanto más larga, mayor es la susceptibilidad al ruido.
- Blindaje (no necesario)
- Más delgado que AWG28

6.2.2 Sensor AD8232

Es una tarjeta de bajo costo usada para medir la actividad eléctrica del corazón. Esta actividad eléctrica puede ser registrada como un electrocardiograma ECG o una salida como una lectura analógica. Proporciona nueve conexiones desde el circuito en el que puedes soldar pines, cables u otras conexiones. Las salidas SDN, LO+, LO-, OUTPUT, 3.3V y GND proporcionan los pines esenciales para operar el monitor con un Arduino u otra tarjeta de

desarrollo. También proporcionan en esta tarjeta los pines RA (Right Arm), LA (Left Arm) y RL (Right Leg) para conectarlos y usar tus propios sensores. Adicionalmente hay un LED indicador que emite pulsos luminosos al ritmo del bombeo del corazón. Los sensores biomédicos -electrodos- y los cables para electrodos se requieren para usar el monitor.

Especificaciones: Anexo C (Hoja de especificaciones)

Modelo: Modulo AD8232 ECG Pulso Cardíaco.
Dimensiones: 28mm x 35mm.
Voltaje de operación: 3.3 V.
Bajo consumo de corriente: 170 uA.
Salida de tipo analógica.

Rechazo de ruido a 60Hz: 80dB.
Configuraciones: 2 o 3 electrodos. (Para este trabajo se hará uso de 3 electrodos)
Ganancia elevada ($G=100$), con bloqueo de corriente DC.

6.3 Procedimientos:

Para realizar el procedimiento se definieron los siguientes pasos para resumir el proceso de la implementación del dispositivo para la obtención de datos a la generación de la base de datos:



Figura 4 Diagrama de bloques solución propuesta.

1. Se realizó un proceso de implementación en los sensores y dispositivos, teniendo como propósito acondicionar las señales (en caso de ser necesario acoplar impedancias, reducir el ruido o aplicar filtros). Generó una tabla especificando las diferencias entre la base de datos del MIT y de nuestra base de datos. Condicionando que fueron avaladas por profesionales MIST y base de datos publicas etiquetas por profesionales. Comprobó que tienen condiciones similares para todo el conjunto de la muestra estadística (lo que se considera como el siguiente objetivo a lograr). Igualmente, para la obtención de las muestras se siguió todos los lineamientos del Tratado de Helsinki.

2. Adquirió señales EEG y ECG encontrando la combinación que mejor se ajustó al entrenamiento de la red neuronal. Obteniendo una red ya entrenada con datos de calidad que permita hacer predicciones o clasificaciones acertadas. Desarrollo de un preprocesamiento de datos para facilitar la construcción de una base de datos, para mejorar el entrenamiento e implementar un modelo efectivo de clasificación.

3. Extrajo patrones de las señales e imágenes que nos sirvan para categorizar el tipo de nivel de estrés actual de una persona. Apoyando de una base de datos en de Sistema de Gestión de Bases de Datos Relacionales (por sus siglas SQL) que permitió almacenar toda la información deseada en bases de datos relacionales. Una vez entrenada las redes

neuronales y los algoritmos para el análisis de las señales fisiológicas. Llevo a cabo un proceso para detectar tendencias y patrones del conjunto de datos. Delimitado por un algoritmo que basándonos en el modelo entrenado fue un indicador de la frecuencia, duración y nivel de estímulo que le provoca estrés.

4. Verifico los resultados obtenidos se utilizó para pruebas psicológicas avaladas, verificación de especialistas de los datos brindados. En casos que se requiera test de cortisol, análisis de señal EMG trapecio, dilatación de la pupila o biomarcadores para corroborar que el modelo de aprendizaje automático funciona como clasificador, predictor y para monitoreo de tendencias de nivel de estrés. Debido a que en las anteriores pruebas ya se comprobó su eficacia para identificar el nivel de estrés. Buscó obtener la consulta de instituciones de esta especialidad con el fin de obtener una retroalimentación.

6.3.1 Acondicionamiento de sensores ECG y EEG

Señal ECG

El siguiente paso realizado la conexión de Arduino UNO con el sensor ECG como se muestra en la figura 5:

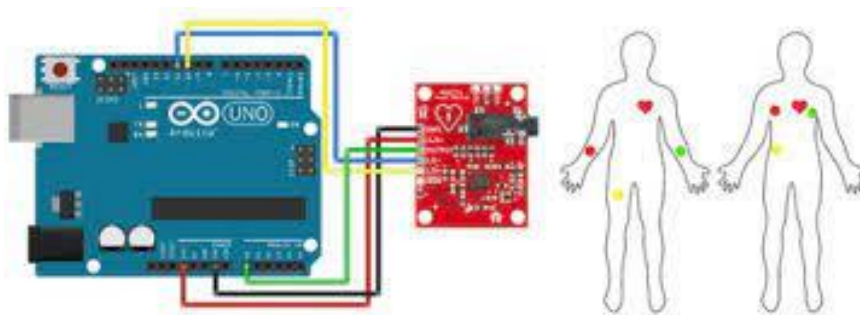


Figura 5 Diagrama conexión Arduino UNO y sensor ECG. Recuperado [31]

El Módulo AD8232 ECG cuenta con 6 pines, para conectarlo al Arduino UNO solo se necesita 5 pines. Continuando con el acondicionamiento de los dispositivos, lo principal es disponer de un electrodo y el chip. Para el acondicionamiento de la señal ECG delimitó la especificación del módulo AD8232, definiendo a este como un integrado con amplificadores de señal y filtros de ruido especialmente calibrados para las señales ECG, evitó la parte de crear un sistema de acondicionamiento físico. Por defecto el módulo es capaz de suprimir el ruido de 60 Hz, generado por la electricidad doméstica. La salida del módulo es de tipo analógico, fue necesario soldar los pines y conectar el módulo a un microcontrolador como Arduino como se muestra en la figura anteriormente en la figura 5. Dentro del programa se realizó la conversión analógica a digital y se observó el ECG en el comando "Plotter" de Arduino al conectar los electrodos como se muestra siendo la figura 5, en la que se ilustra la forma en que se colocaron los electrodos para adquirir las señales.

Implementación de señal EEG y señal ECG

Se desarrollo los dispositivos necesarios para acondicionar las señales: Sensor AD8232. Un integrado con amplificadores de señal y filtros de ruido especialmente calibrados para las señales (suprime el ruido de 60 Hz, por lo que no se necesitó método adicional de acondicionamiento). Se adquirió e implementó el hardware de MindFlex, contenía un software para la adquisición de datos de la banda de frecuencia EEG. El microcontrolador en la diadema analiza los datos del chip EEG y envía actualizaciones de manera inalámbrica a una estación base, en este caso se conjuntó con la tarjeta de desarrollo Raspberry Pi-Mindflex, proporcionaba ocho valores que representan la cantidad de actividad eléctrica a diferentes frecuencias. Estos datos se filtraban y simplificaban en gran medida, por lo que cuando un EEG de grado médico convencional le da valores de voltaje absolutos para cada banda, NeuroSky.

En su lugar brindo mediciones relativas que no se asignan fácilmente a unidades del mundo real. Los datos que entrega este hardware fue la señal aplicando una transformada rápida de Fourier (por sus siglas) para las bandas de potencia EEG que entregaron, este proceso es el que fue programado por defecto en el dispositivo. Así que no se necesitó realizar algún filtro para este, se usó filtros y limpieza de datos que tenía como algoritmos definidos por el creador del dispositivo (aunque se implementó un filtro butterworth pasa bajas en Python porque no fue suficiente). El preprocesamiento de datos, adjunto técnicas para eliminar valores atípicos; ya que en ocasiones se perdieron datos [18].

Continuando para decodificar el chip de Mindflex se cambió el protocolo de comunicaciones para el chip ThinkGear al modo 0x02 enviando 0x02 a su línea de recepción en serie. Asegurando de que la línea de recepción del chip fue conectada a la línea de transmisión de la radio Bluetooth y envíe 0x02. El único problema es que al enviar cosas a través de la serie RS232 o su variante TTL, no se podía controlar cada bit. La marca y los bits de parada estaban fuera de nuestro control [18] Por lo que como alternativa se abrió el circuito y se colocaron dos cables en RX y la tierra para evitar problemas de comunicación.

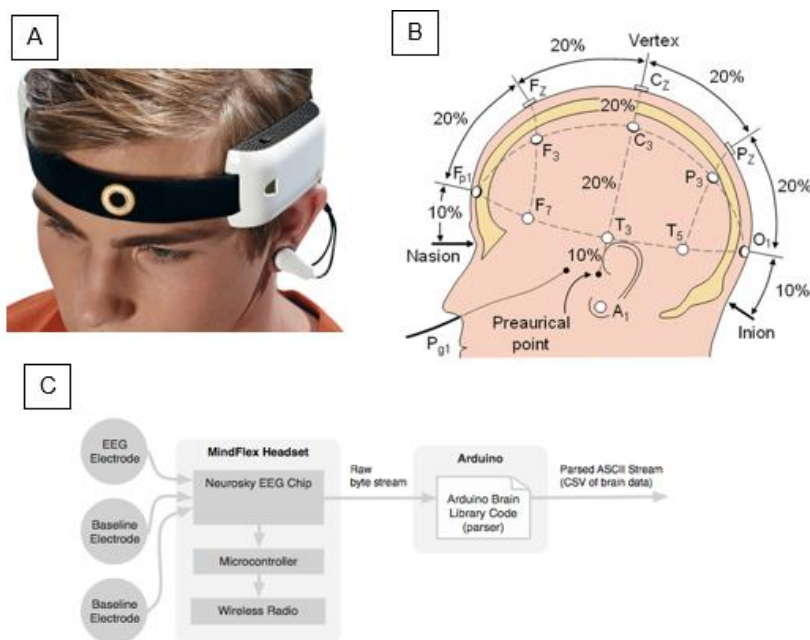


Figura 6 Diagrama Arduino y Mindflex. Recuperado [29][32][33]

En la primera figura 6 A se observa el dispositivo en su presentación comercial, en este se observa sus características físicas y la forma en que este dispositivo se tenía que colocar para la lectura de datos. Recurriendo a la segunda figura 6 B se observa los puntos que son colocados el electrodo, son los puntos A1 y A2 que sirvió como referencia. Y existe otro electrodo colocado en el punto Fp1 y Fp2 (todo esto siguiendo los estándares de Sistema internacional 10-20). Y por último, en la parte de abajo la figura 6 C muestra cómo se implementa la adquisición de datos EEG desde un punto de vista general, los materiales que se requieren y el proceso explicado. Se realizó un programa en Python recopilando estas variables, recalcando que en el proyecto final no se visualizará como muestra la figura 7. Solamente se almacenarán los datos y las variables requeridas para almacenarlo en una base de datos SQL. De manera que las señales se convirtieron en datos resultantes al aplicar el método FTT a las señales EEG y la señal ECG fue la señal original acotada por las variables estadísticas generadas y definidas como predictores de importancia.

Nota: Al tomar la señal ECG y EEG se utiliza 84 % de la memoria RAM de las Raspberry Pi, motivo por el que en el momento de la interfaz no se contempla mostrar las señales adquiridas en el instante en que se adquieren por la falta de recursos computacionales.

6.4 Análisis teórico:

6.4.1 Descripción de variables análisis EEG

Podemos ver que existen datos atípicos en las ocho variables analizadas, recordemos que estas se refieren a ondas cerebrales por lo que el incremento de Hertz significa una tensión de alguna forma o relajación extrema.

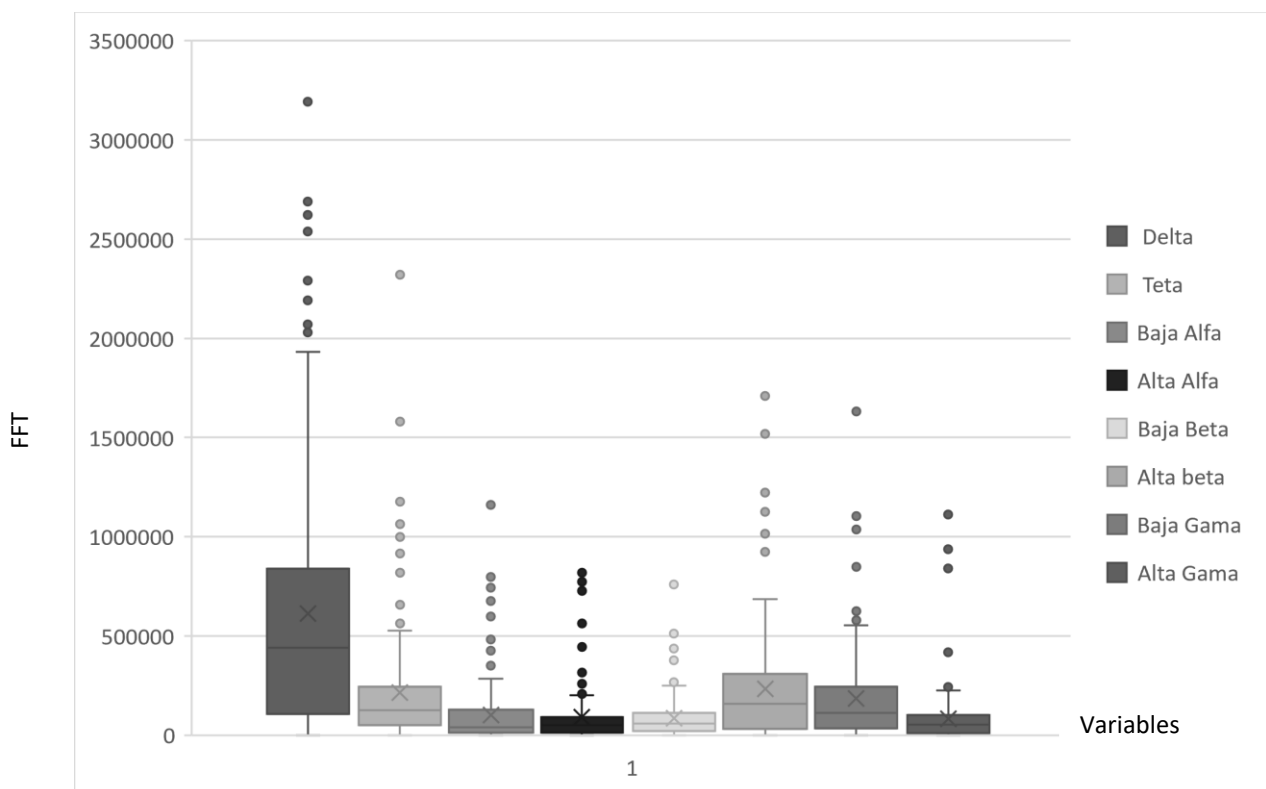


Figura 7 Diagrama de caja EEG

Con la figura 8 se observa una existencia de datos atípicos en todas las ondas, pero estos no fueron excluidos del modelo de clasificación debido a que antes de eso se aplicó un preprocesamiento.

6.4.2 Preprocesamiento de datos y adquisición de datos



Figura 8 Diagrama preprocesamiento

Lo primero que se realizó fue eliminar datos que se encontraban a alejados del primero y último percentil. También se aplicó una discriminación de datos por distancia de Mahalanobis para eliminar datos fuera de esa distancia. Y para entrenar nuestro modelo se convirtieron los datos necesarios en dummy y se normalizaron los datos, véase la figura 9.

Ese accionar se tiene que establecer para la base de datos, debido a que se encontraba con valores atípicos y no se podrá realizar todas las variables. Por recursos computacionales se tiene que optimizar el número de variables en función a reducirla y gastar menos memoria para manejar un control de datos que no ocasionen un problema en el sistema general.

6.4.3 Correlación entre variables

En el apartado anterior se expusieron los posibles padecimientos o estados y/o sentimientos derivados de nivel bajos o altos de cada onda analizada, en este apartado se estudiará la relación entre estas.

Tabla 3 Correlaciones EEG

Columna1	Delta	Teta	Baja Alfa	Alta Alfa	Baja Beta	Alta beta	Baja Gama	Alta Gama
Delta	1	0.39	0.3	0.3	0.42	0.4	0.38	0.37
Teta	0.39	1	0.5	0.52	0.55	0.52	0.45	0.49
Baja Alfa	0.3	0.5	1	0.59	0.53	0.37	0.31	0.47
Alta Alfa	0.3	0.52	0.59	1	0.59	0.37	0.33	0.6
Baja beta	0.42	0.55	0.53	0.59	1	0.5	0.45	0.66
Alta beta	0.4	0.52	0.37	0.37	0.5	1	0.77	0.609
Baja Gama	0.38	0.45	0.31	0.33	0.45	0.77	1	0.566
Alta Gama	0.37	0.49	0.47	0.6	0.66	0.609	0.566	1

En la Tabla 3 podemos observar que la relación entre las ondas cerebrales estudiadas es positiva entre las variables, pero esta tiende a ser baja, lo que nos permite establecer que el conjunto de variables no tiende a moverse en la misma dirección.

Tabla 4 Correlación r y absoluta

Variable Uno ▼	Variable Dos ▼	r ▼	Absoluto® ▼
Baja Beta	Alta Alfa	0.597	0.597
Alta Alfa	Baja Beta	0.597	0.597
Baja Alfa	Alta Alfa	0.596	0.596
Alta Alfa	Baja Alfa	0.596	0.596
Baja Beta	Teta	0.557	0.557
Teta	Baja Beta	0.557	0.557
Baja Beta	Baja Alfa	0.531	0.531
Baja Alfa	Baja Beta	0.531	0.531
Teta	Alta Alfa	0.527	0.527
Alta Alfa	Teta	0.527	0.527

Se observa en la Tabla 4 como las señales que tienen más correlación entre ellas, tomando en cuenta varía la señal que tiene más peso en el modelo varía dependiendo el dispositivo y el método de clasificación. A causa de lo anterior se realizó el siguiente análisis de factores y PCA:

Columna1 ▼	PA1 ▼	PA2 ▼	h2 ▼	u2 ▼	com ▼
Teta	0.58	0.39	0.48	0.52	1.7
Baja Alfa	0.69	0.19	0.51	0.49	1.1
Alta Beta	0.81	0.17	0.68	0.32	1.1
Baja beta	0.69	0.37	0.62	0.38	1.5
Alta beta	0.3	0.85	0.82	0.18	1.2
Baja Gama	0.23	0.83	0.74	0.26	1.2
Alta Gama	0.61	0.5	0.62	0.38	1.9

Figura 8 Análisis factorial EEG.

Partiendo del análisis de componentes principales comenzamos con un análisis con dos factores, sin embargo, notamos que la onda delta, la encargada del sueño, podría verse como un factor por sí mismo. En consecuencia, se busca corroborar con la matriz de relación mostrada en apartados anteriores, en esta podemos ver que la variable delta es una de las menos correlacionadas con el resto de las variables. De esta forma se vuelve a correr el análisis de dos factores, excluyendo la variable delta, dado que se asume como factor propio, de esta forma con una rotación PROMAX se obtienen los resultados de la figura 9.

Finalmente, a pesar de que al desarrollar el análisis de componentes principales y de factores nos arrojan que las variables podrían ser entre dos o tres, podemos decir que no es necesario disminuir el número de variables a estudiar. Cada onda cerebral tiene que aportar al diagnóstico del nivel de estrés detectado en cada individuo, sin embargo, a la onda delta es la de mayor peso en la detección del nivel de estrés. Al implementar la técnica de PCA se concluyó que la técnica sirve para simplificar la complejidad de espacios muestrales con muchas dimensiones a la vez que conserva su información. Pero no permitió un análisis detallado de las variables a través del peso de estas. No se necesita descartar componentes para acercarnos a un 90% de la variable explicativa ya que estos dos componentes representan el total de varianza explicada. Y al crear los componentes principales ninguna se compone en su mayoría por un tipo de onda tiene un valor en el componente semejante. En este caso solo se definió dos componentes como se muestra en la figura 10:

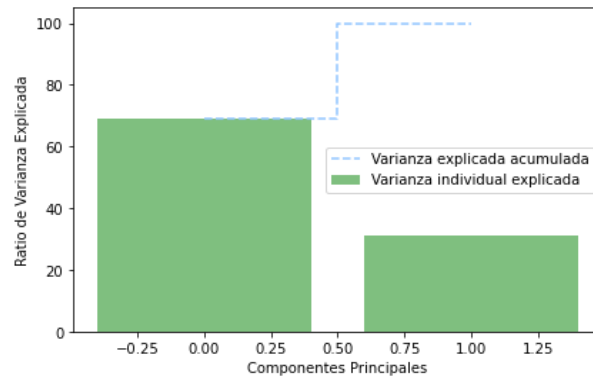


Figura 9 Radio de varianza explicada

A pesar de que al desarrollar el análisis de componentes principales y de factores nos arrojan que las variables podrían ser entre dos o tres, podemos decir que no es necesario disminuir el número de variables a estudiar. Esto afirma que cada onda cerebral tiene que aportar al diagnóstico del nivel de estrés detectado en cada individuo, sin embargo, la onda delta es la de mayor peso en la detección del nivel de estrés. Y en cuanto al análisis de componentes principales, sirvió para simplificar la complejidad de espacios muestrales con muchas dimensiones a la vez que conserva su información. Pero no permitió un análisis detallado de las variables a través del peso de estas.

Mencionado lo anterior se decidió por el análisis de factores realizados tomar delta y alfa para la adquisición de características.

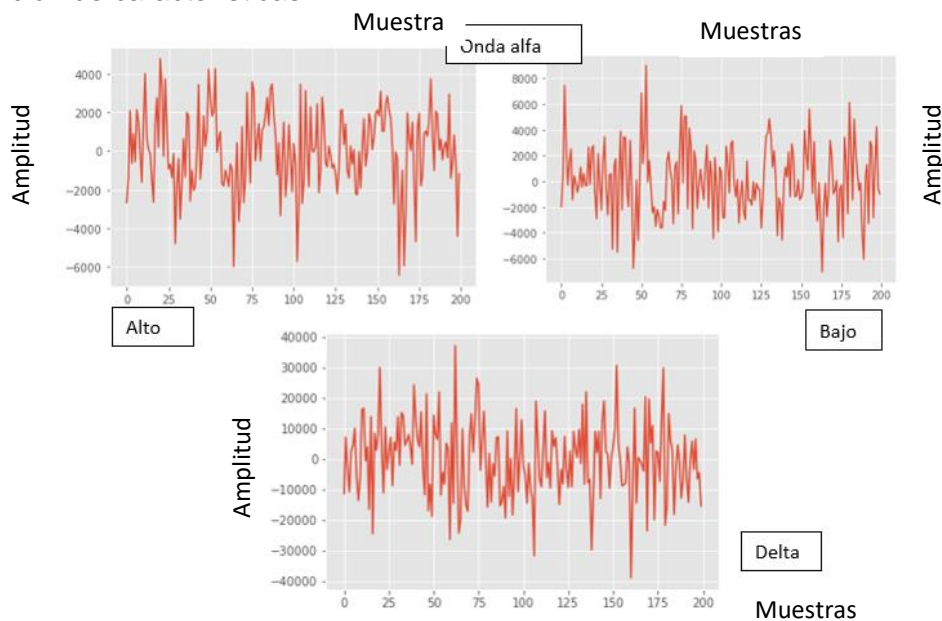


Figura 10 Muestra de señales EEG (alta alfa, baja alfa y delta)

Nota: En la figura no muestra unidades se desconoce la escala de la señal, por lo que solo se puede mencionar que representa muestras en relación con la señal de forma adimensional. Resultado de analizar las señales adquiridas se generó el siguiente vector con las características para su clasificación.

6.4.4 Descripción de variables análisis ECG

El análisis inició creando la base de datos, una vez recopilada se buscó características en las cuales favorecen el aprendizaje supervisado. Se comenzó desarrollando un árbol de decisión para delimitar la importancia de los predictores del modelo. La base de datos inicial contiene una propuesta de análisis de más de 35 características de la señal ECG como estadísticos. Por lo cual la frecuencia de muestreo no resultó un problema porque no afecta de forma crítica a las medidas de tendencia central y dispersión. Debido a lo anterior no es necesario tener un dispositivo con las mismas características de adquisición de datos, ya sea frecuencia o presión [20]. Aunque se debe de establecer como estándar que no existiera offset en la señal y tampoco existiera ruido en ella. La base de datos que se presenta es la recopilación de datos públicos, añadiendo algunas señales adquiridas y quitando algunas variables al aplicar el preprocesamiento en esta.

Tabla 5 Compendio de base de datos recopilada ECG [24]

	MEAN_RR	MEDIAN_RR	SDRR	RMSSD	SDSD	SDRR_RMSSD	HR	pNN25	pNN50	SD1	...	LF_NU	HF	HF_PCT	
0	885.157845	853.763730	140.972741	15.554505	15.553371	9.063146	69.499952	11.133333	0.533333	11.001565	...	98.485263	15.522603	0.421047	1
1	939.425371	948.357865	81.317742	12.964439	12.964195	6.272369	64.363150	5.600000	0.000000	9.170129	...	99.695397	2.108525	0.070133	0
2	898.186047	907.006860	84.497236	16.305279	16.305274	5.182201	67.450066	13.066667	0.200000	11.533417	...	98.950472	13.769729	0.512671	1
3	881.757865	893.460030	90.370537	15.720468	15.720068	5.748591	68.809562	11.800000	0.133333	11.119476	...	98.224706	18.181913	0.529387	1
4	809.625331	811.184865	62.766242	19.213819	19.213657	3.266724	74.565728	20.200000	0.200000	13.590641	...	96.720007	48.215822	1.839473	3
5	923.283866	617.794160	517.536544	9.965976	9.933933	51.930344	81.342254	1.200000	0.600000	7.026695	...	93.200171	11.027460	0.318490	6
6	973.252908	964.650020	82.405179	10.644196	10.643638	7.741794	62.095066	2.000000	0.000000	7.528700	...	99.692575	1.489796	0.119828	0
7	715.914682	679.499395	131.477151	9.477727	9.477717	13.872224	85.857703	2.533333	0.200000	6.703994	...	93.805918	28.913453	1.501528	6
8	814.257021	827.522830	87.014459	14.632232	14.631275	5.946766	74.588857	7.733333	0.800000	10.349326	...	97.973018	20.757787	0.659188	2
9	959.694591	957.895600	54.904529	12.015400	12.015343	4.569513	62.726998	3.266667	0.200000	8.498966	...	99.509898	2.572459	0.150800	0

10 rows × 35 columns

El proceso de preprocesamiento consistió en eliminar datos arriba del 97 % y 3 % (esto se realizó debido a que los datos representaban un fallo en la medición de la señal o calidad, y estos tendían a ir a los extremos. Es decir, la variable se convertía en unos de los valores más grandes o se volvía una señal inexistente, representado por un cero) de las muestras y también utilizó distancia de Mahalanobis para eliminar datos, se convirtieron los datos necesarios en dummy además de normalizar datos. Se utilizó árbol de decisión para ver que muestras tenían más peso, con la diferencia del primer reporte que con solo 6 variables representan el 71.2 % de los predictores del modelo.

Por lo que se concentró en usar esos 6 predictores. Debido a que el modelo representa un error (rmse) de prueba es: 0.0257. Por último, en esa base de datos se verificó que no existiera problema de colinealidad, pero este se descartó, ya que ninguno superó un valor de cinco.

Resultando como base final para análisis de ECG las siguientes variables:

Tabla 6 Base de datos con los seis predictores más importantes

	HR	MEDIAN_RR	LF	MEDIAN_REL_RR	sampen	pNN50	condition_no stress
0	84.121868	727.267280	615.914573	-0.000951	2.097342	0.000000	1
1	71.478642	844.407930	1566.866135	-0.001543	2.217275	0.200000	0
2	63.874293	966.671125	2074.868884	-0.002000	2.217136	1.800000	1
3	74.330531	842.485905	505.886664	0.000445	2.106863	0.533333	1
4	82.092049	747.941620	524.203971	-0.000170	1.912191	0.666667	0
5	72.562697	836.596795	586.603990	0.000707	2.177956	0.266667	0

Donde se analiza una única señal, la cual como se mencionó con anterioridad esta tuvo un preprocesado de filtro Butterworth pasa bajas y quitar la oscilación de las señales:

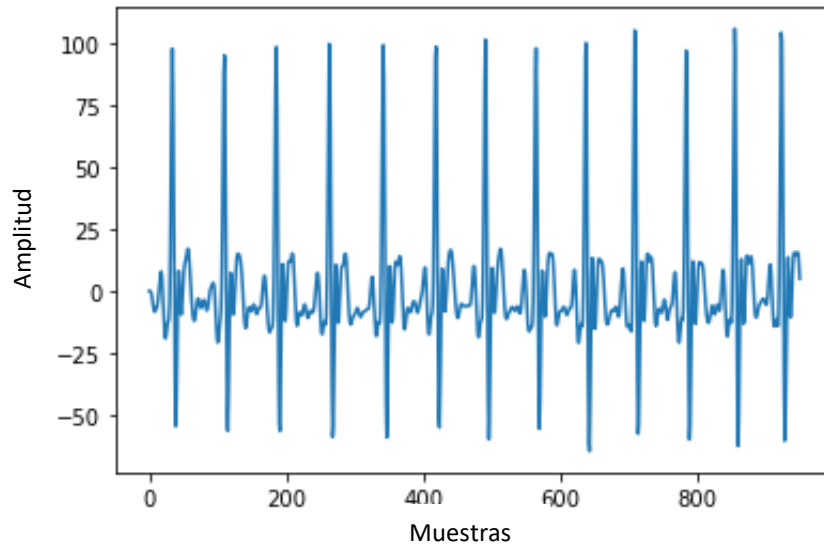


Figura 11 Muestra de señales ECG

Y se obtienen las siguientes características:

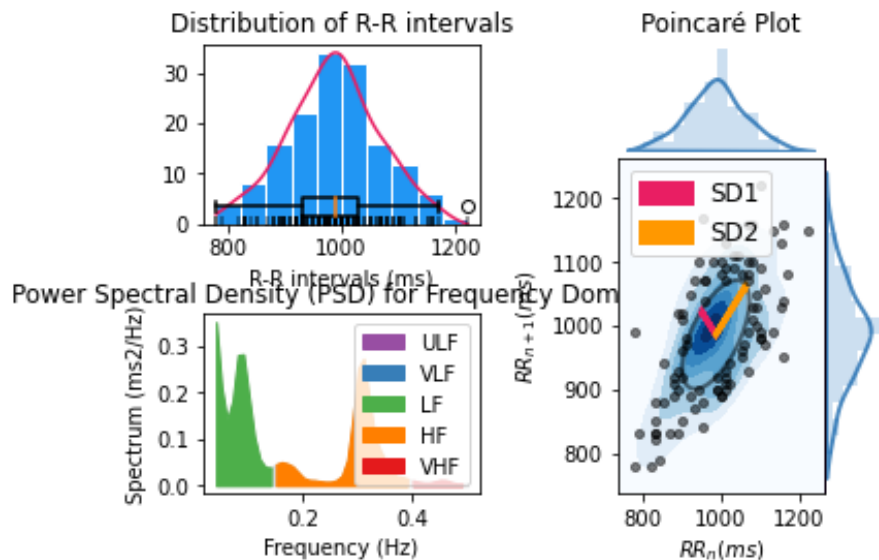


Figura 12 Análisis de intervalos y distribución RR

Para llegar a ese resultado se realizó un análisis exploratorio de los datos, el primer paso que se realizó fue realizar un árbol de decisión de ramificaciones que no son extensas y tampoco profundas donde se observaron los parámetros más importantes para delimitar si existe estrés.

Profundidad del árbol: 3

Número de nodos terminales: 8

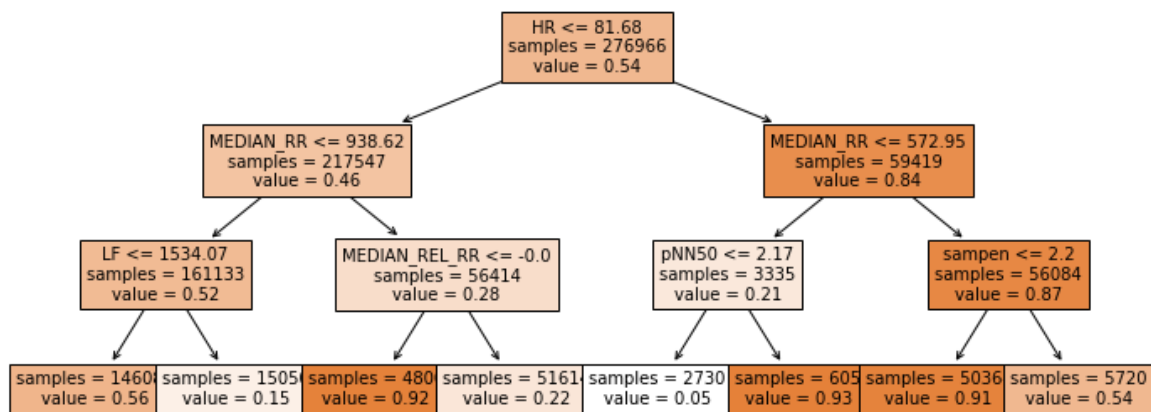


Figura 13 Diagrama de árbol de decisión ECG

Este árbol de decisión con un 88.5 % de efectividad fueron delimitados los predictores más importantes en donde se destacaron seis que serán las variables que se conservará (HR, Mediana RR, LF, Mediana, Sampen y pNN50) y que se implementó el programa para calcular estas. Los resultados son congruentes por lo que el entramiento para implementar el árbol de decisión tendrá las especificaciones antes mostradas.

Presenta una congruencia con nuestros predictores de mayor importancia arrojados por el modelo, mencionando que este modelo se buscó una relación entre baja profundidad y ramificaciones del árbol para evitar trabajar con una gran cantidad de variables. Finalmente, la cuantificación del nivel de importancia de los predictores es la siguiente:

	predictor	importancia
6	HR	0.414069
1	MEDIAN_RR	0.241789
23	LF	0.140929
14	MEDIAN_REL_RR	0.134902
32	sampen	0.044416
8	pNN50	0.023895
0	MEAN_RR	0.000000
25	LF_NU	0.000000

Figura 14 Base de datos con los seis predictores más importantes representados en porcentaje ECG.

Ese fue el resultado anterior a implementar un árbol de decisión con más profundidad y número de terminales, lo que permitió disminuir el tiempo de entrenamiento y facilitar la interpretación de los datos para la toma de decisión.

Se realizó para reducir la varianza del modelo y mejorar la capacidad de predicción, se poda el árbol. Como se mencionó anteriormente, el proceso de poda intenta encontrar el árbol más simple (tamaño mínimo) que logra el mejor resultado de predicción. Para aplicar el proceso de poda, se debe especificar el parámetro, el parámetro de la frecuencia de banda de la onda Alfa de la señal EEG, que determina el grado de penalización por complejidad. Cuanto mayor sea el valor, más agresiva será la poda y menor será el tamaño del árbol resultante como lo fue el anterior presentado. Dado que el valor óptimo de la onda Alfa no se puede conocer de antemano, se utiliza la validación cruzada para identificarlo.

Aunque hay otras formas de identificar el mejor árbol, como identificar los valores de profundidad máxima y el mínimo de muestras mediante validación cruzada, la poda puede producir mejores resultados porque comienza la siguiente división a partir de él, permite mantener los predictores más determinantes. Y el comportamiento de los predictores iniciales antes de aplicar la técnica de podado del árbol eran los siguientes:

Tabla 7 Base de datos con los seis predictores más importantes representados en porcentaje antes del pruning ECG.

	predictor	importancia
1	MEDIAN_RR	0.131138
14	MEDIAN_REL_RR	0.087409
7	pNN25	0.059714
0	MEAN_RR	0.058835
29	TP	0.049690
33	higuci	0.049070
5	SDRR_RMSSD	0.047520
18	SDRR_RMSSD_REL_RR	0.044749
8	pNN50	0.043967
21	VLF	0.042704

De esta forma finaliza el análisis descriptivo para realizar los modelos de clasificación, con los predictores definidos, así como delimitado el rango de error.

6.4.5 Técnica no supervisada clasificación

Lo primero que se realizó como se comentó el uso de un clustering para corroborar la hipótesis si se podía sub categorizar el nivel de estrés con relaciones entre ondas EEG. Por lo cual se comenzó con KNN y distancia Euclidiana representada en estos dos dendrogramas. Donde sé al proporcionar la información necesaria para combinar o separar categorías de datos, los diagramas de árbol se pueden utilizar para reducir la clasificación estadística errónea del análisis. Si las clases en su análisis son estadísticamente muy cercanas, es decir, es difícil distinguir las dos clases según sus estadísticas, puede causar una clasificación errónea. En este caso, se consideró combinar estas clases, lo cual como se observa esta separación no es trivial, por lo cual se usó KNN. Obteniendo lo siguiente:

high_beta	low_gamma	high_gamma	Cluster	Delta_Low	Betha
204057	262653	64201	2		
461158	193445	88907	6		
181713	328328	71307	2		
223951	362543	56752	1		
131624	185840	58791	3		
...		
11373	13563	10029	1		
52574	150696	83628	4		
24708	12588	15166	1		
30419	20595	10593	6		
88830	81687	92653	4		

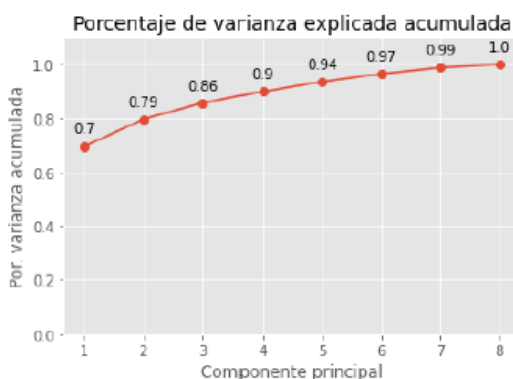


Figura 15 Clúster análisis señal EECG

Su función es obtener grupos de objetos de tal manera que, por un lado, los objetos pertenecientes a un mismo grupo sean muy similares entre sí, es decir, el estrés, y, por otro lado, los objetos pertenecientes a diferentes grupos tienen comportamientos diferentes respecto a las varianzas. Aunque en este caso de estrés no se encuentra una ventaja en

realizar esta subclasificación, solo sería agregar otra variable para el aprendizaje. Ya que como se consideró que lo que se busca es tener resultados semejantes a los concluidos en este trabajo con las diferencias antes estipuladas [21]. Las señales de sensor obtenidas se dividieron en ventanas de intervalos de 5, 10 y 15 s para crear tres conjuntos de datos diferentes y se definieron 112 características diferentes a partir de los datos sin procesar (como se buscará manejar en nuestro sistema de clasificación multi clase del nivel de estrés). Para obtener subconjuntos de características más efectivos, estas características se clasificaron utilizando el algoritmo de selección de características Gain Ratio, lo cual se aplicará de la misma forma en nuestro proyecto si nuestros datos lo permiten. Después, Los comportamientos de estímulos de estrés se clasifican mediante los métodos de árboles de decisión, redes bayesianas o k-vecino más cercano para realizar la predicción de este estímulo. Como resultado buscamos tener una efectividad semejante a la que se obtuvieron en los resultados de clasificación de precisión del 74,26 %, 67,86 % y 87,56 %, del proyecto citado [22]. La cual no se alcanzó debido a que nuestra base de datos tenía mayor variedad en la muestra estadística.

6.4.6 Técnica supervisada algoritmo de clasificación

Se logró un acercamiento de qué técnica se usará para lograr la clasificación de este. Ya que aún no se han probado algunas arquitecturas de redes neuronales o por ejemplo regresión. En primera instancia se tienen resultados similares a los establecidos en la bibliografía donde indican que no es una buena opción combinar datos de estos dos para el entrenamiento, por lo que se buscará un modelo de probabilidad o estadístico para verificar si estos eventos son independientes o dependientes entre sí. Para eso se verá desde el punto de vista biológico y respaldado desde la parte matemática. Añadiendo al trabajo me gustaría denotar que aumentar el número de variables al entrenamiento aumenta la efectividad, pero el tiempo en entrenar aumenta, por lo que se tiene que evaluar el perder el número de aciertos y los recursos computacionales que tiene la tarjeta de desarrollo.

Por lo que el segundo paso se denotará por concentrarse en la red neuronal para la señal EEG, como se especifica en el objetivo específico en este trabajo. Solo que se realizará un modelo donde en dado caso que no aumente el número de aciertos baje el peso de esta señal para la evaluación final, si la persona tiene estrés o no [23]. También se presentó como avance el protocolo de cómo se adquirirán las variables para hacer una base de datos que después de cierto tiempo se personalice a los estándares de la persona que se evalúa su nivel de estrés. Así como la implementación del software. Se muestra la comparativa de los diferentes métodos de implementados

Tabla 8 Resultado de análisis supervisado

	Base de datos por separado base de datos testing	Base de datos ECG y EEG base de datos testing
Vanilla neural network	EEG: 52.3% ECG: 50.2%	EEG y EEG: 40.14%
Perceptron multicapa con MLPClassifier	EEG: 56.3% ECG: 55.4%	EEG y EEG: 49.2%
Backpropagation	ECG:62.1% EEG:55.8%	EEG y EEG: 25.9%
Visualización de pesos MLP en MNIST	ECG:65.1% EEG:64.3%	EEG y EEG: 53.7%
Árbol de de decisión	ECG:90.1% EEG:88.5%	EEG y EEG: 79.6%
Random forest classifier	ECG:82.3% EEG:80.9%	EEG y EEG: 79.3%

usará para lograr la clasificación de este. Ya que aún no se han probado algunas arquitecturas de redes neuronales o por ejemplo regresión. En primera instancia se tienen resultados similares a los establecidos en la bibliografía donde indican que no es una buena opción combinar datos de estos dos para el entrenamiento. Por consecuencia, se buscará un modelo de probabilidad o estadístico para verificar si estos eventos son independientes o dependientes entre sí. Para eso se verá desde el punto de vista biológico y respaldado desde la parte matemática. Añadiendo al trabajo me gustaría denotar que aumentar el número de variables al entrenamiento aumenta la efectividad, pero el tiempo en entrenar aumenta, por lo que se tiene que evaluar el perder el número de aciertos y los recursos computacionales que tiene la tarjeta de desarrollo.

Tabla 9 Red neuronal autoencoder

Datos:	Porcentaje	Total
Entrenar	80%	716 donde $y=0$
Prueba	20 %	370 donde 120 $y=1$

Matriz confusión prueba	0	1
0	142	45
1	35	85

Características:

Exactitud: 73.36 %	Precisión: 65.38%	Exhaustividad: 70.83%	Especificidad 70.83%
Puntaje F1: 68%		Falso Positivos: 24.06 %	Falso Negativos: 29.16 %

De modo que el impedimento era el comportamiento que a mayor efectividad aumentaba el resultado de falsos positivos,

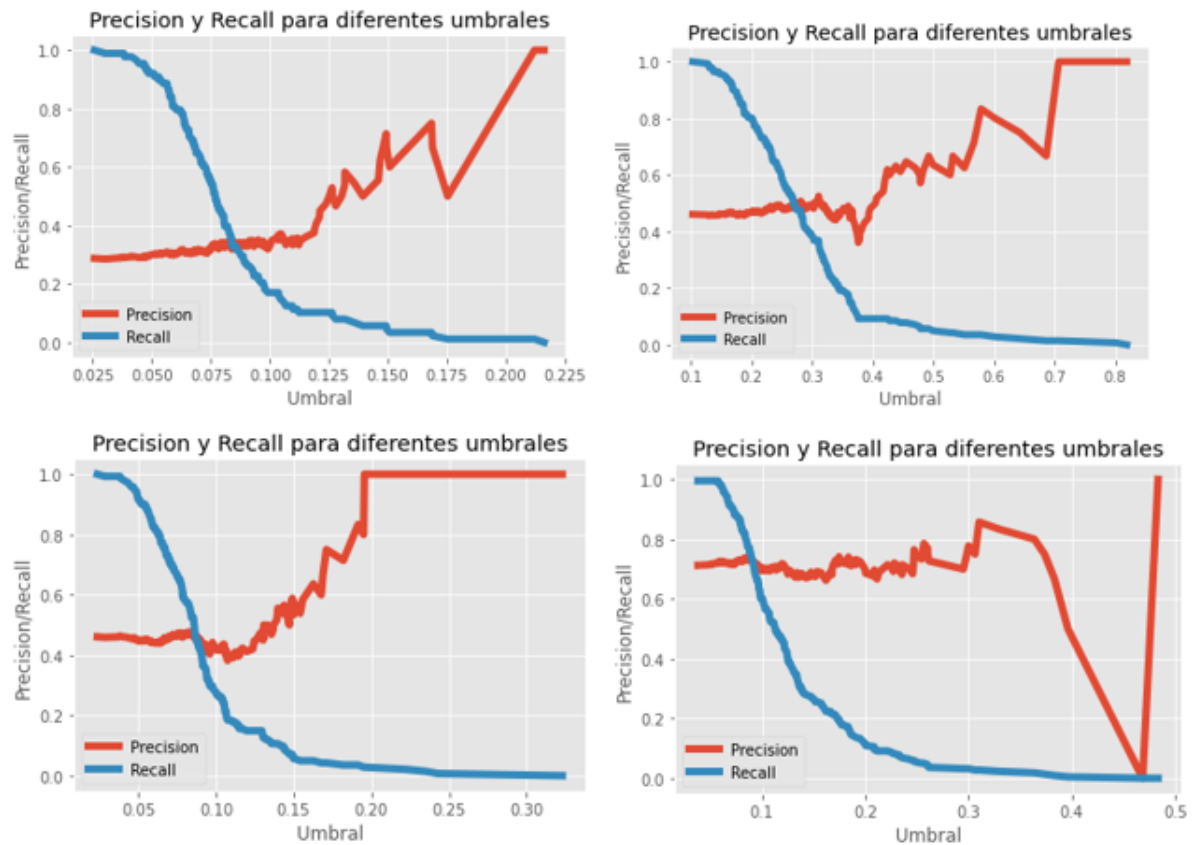


Figura 16 Relación de precisión y exhaustividad

La desventaja es su alto riesgo de sobre entrenamiento, por lo que se utiliza pocas épocas y se verificó la evolución del error en su conjunto de validación. Se eligió tener una alta especificidad, ya que el modelo preferirá que se detecten mayormente falsos negativos y se detecten todas las señales que muestren estrés. Es decir, no se desean falsos positivos y lo que es tolerable en nuestro modelo es que se detecten la totalidad de señales que tengan características para ser delimitadas como estrés.

6.4.7 Muestreo EEP 10

Se realizó la prueba EEP descrita con anterioridad y se digitalizó la respuesta con el fin de poder general la siguiente tabla con la recopilación de estos con el fin de analizar de forma factorial lo más importante e implementarlo en el modelo de clasificación.

Tabla 10 Aplicación EPP 10 muestra 20 personas

Preguntas	Res_ID1	Res_ID2	Res_ID3	Res_ID4	Res_ID5	Res_ID6	Res_ID7	Res_ID8	Res_ID9	Res_ID10
id	1	2	3	4	5	6	7	8	9	10
Afectado inesperadamente	2	3	2	2	2	2	3	4	1	2
Incapaz de controlar cosas	2	3	1	3	3	3	2	3	2	3
Nervioso o estresado	1	1	3	4	2	3	3	2	2	4
Manejar problemas	2	1	2	2	1	3	3	2	3	3
Las cosas van bien	3	3	1	1	2	2	1	3	2	2
No podía afrontar	2	3	1	3	3	2	1	3	1	2
Control de dificultades	2	2	2	2	4	2	3	3	3	2
Todo bajo control	1	2	3	2	2	2	2	2	4	3
Enfado	1	2	2	2	3	1	2	2	3	2
Dificultades se acumulan	1	1	1	2	1	1	2	2	2	1

Realizado el muestreo y la digitalización de resultados se creó un reporte que muestre de manera grafico el resultado de la aplicación de esta escala. Véase el resultado en la siguiente figura:



Figura 17 Aplicación EPP 10 presentación resultados

La prueba de EPP 10 se aplica en cada adquisición de datos, se ha realizado el protocolo tres veces por persona en días distintos. Con el fin de encontrar un estado basal y unas sus características de las señales y por supuesto el estrés percibido.

6.4.8 Diseño de base de datos SQL

Se diseñó la base de datos, aunque esta se puede modificar debido a las variables de cada clase que son varias. Añadiendo que en su mayoría son datos int() y char(). Y se usará como llave primaria la entrada Id.

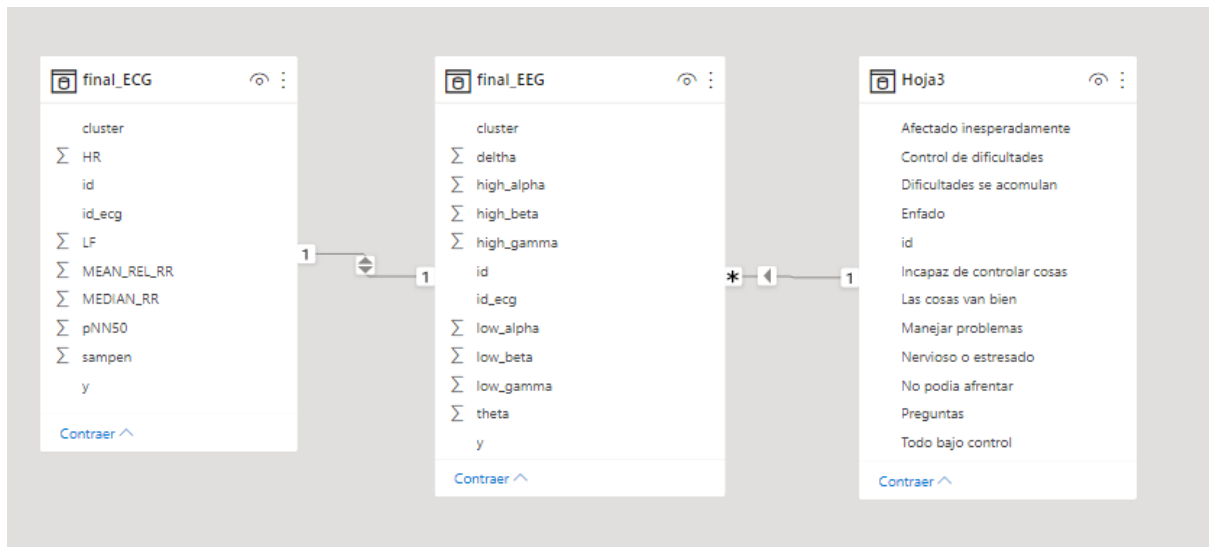


Figura 18 Diseño base de datos SQL

Para conectar la base de datos SQL se hizo uso de la lectura de un archivo de datos en Python y se asignó a una variable, usando la memoria de tu computadora para guardar esta variable. Acceder a datos almacenados en una base de datos como SQL no solo es más eficiente, sino que también te permite extraer e importar todo o partes de la base de datos. Y el módulo elegido para conectar Python y SQL es el módulo `sqlite3` el cual proporciona una interfaz sencilla para interactuar con bases de datos de este tipo. Primeramente, se crea un objeto utilizando `sqlite3.connect()`, esto abre la puerta a la base de datos. Mientras la conexión esté abierta cualquier interacción con la base de datos se creó un objeto cursor, con el comando `cursor()`. Luego el cursor está preparado para realizar todo tipo de operaciones con el comando `execute()`. Y para cerrar la conexión usando el comando `close()`.

6.4.9 Diseño de esquema general del programa

A continuación, se presenta el diagrama de flujo del programa realizado en Python para general la estructura general del programa:

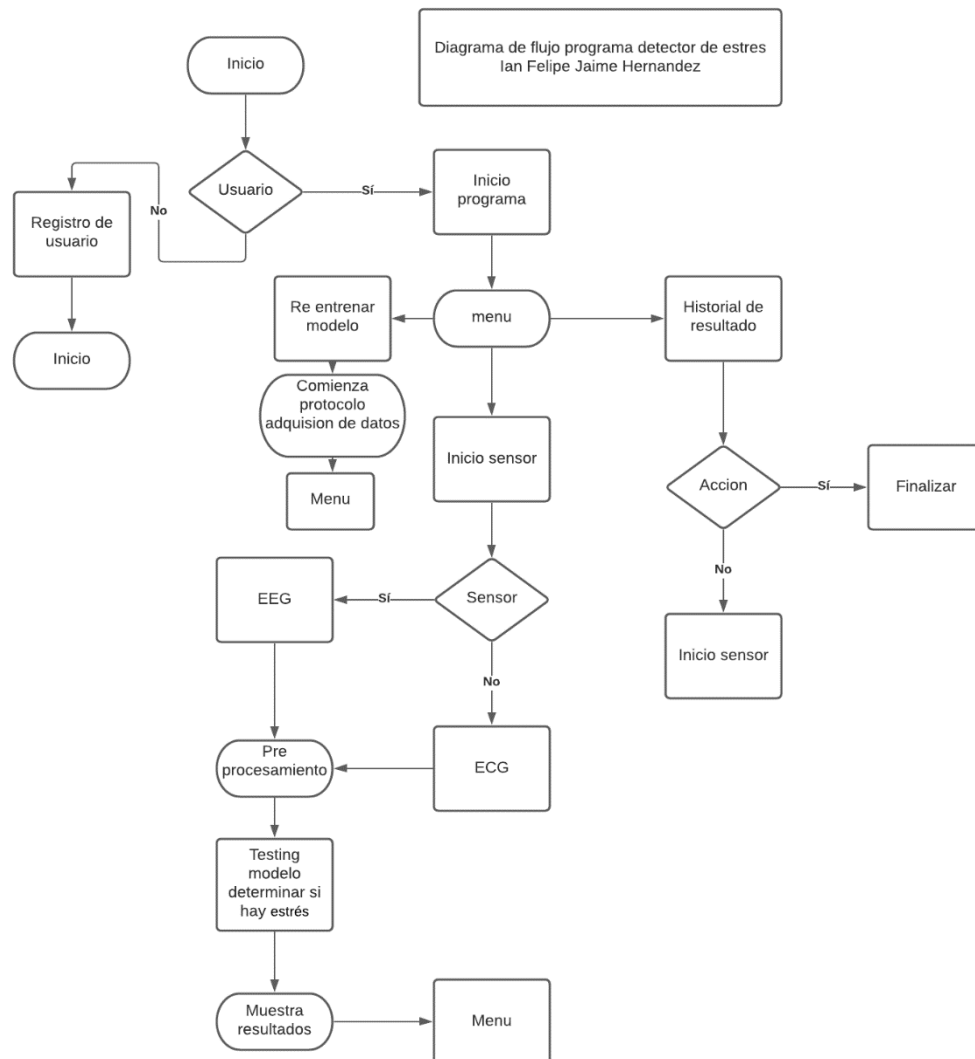


Figura 19 Diagrama de flujo programa general

Con esta herramienta presentada se utilizó a lo largo del proyecto para representar la secuencia de las actividades en un proceso. De modo que con este se pueda usar para desarrollar y mejorar la presentación gráfica de un proceso de adquisición y evaluación del estrés, de forma que se pueda identificar el costo de la calidad del programa.

Para la implementación de software es necesario tomar como base el uso de las tecnologías de código abierto en Linux empleando Arduino y Raspberry Pi. Con el fin de lograr establecer una conexión entre estos dispositivos y poder desarrollar una interfaz a través la cual el

usuario pueda medir su nivel de estrés académico. Y se considera la base de datos realizada como un parámetro para comenzar a diseñar nuestra base de datos.

Para poder almacenar datos se elaboró un sistema electrónico de monitoreo de niveles de estrés académico con el diferencial que será una base relacional y se manejan otros tipos de variable como continuas, discretas y cualitativas, nominales. Además de presentar una tabla relacional que nos facilite la consulta de los datos al profesional médico. Y se contempla poder automatizar el proceso de adquisición de datos de manera que nos facilite una mejor calidad de estos para un mejor entrenamiento. Teniendo como semejanza la idea de implementar una base de datos, pero como diferencia de este proyecto se desarrollará una base de datos relacional y normalizada. Con una conexión de Python y SQL para un almacenamiento de datos de forma estructura para conformar la base de datos. En este trabajo no hace uso de técnicas de inteligencia artificial, su propósito principal fue desarrollar una base de datos, por lo que este proyecto permitirá tener una estructura para el diseño de nuestro proyecto. Acorde al tipo de datos que se han delimitado marcando a este punto como el avance en referencia a este proyecto.

Capítulo 7 Resultados

El dispositivo para estimar el nivel de estrés que se desarrolla en este trabajo terminal tiene las siguientes características:

Tabla 11 Características dispositivo EEG

Canales	Tipo de condición	Sujetos	Frecuencia de bandas	Características	Puntos EEG	Tipo de estresor	Región cerebro
3	Estrés/ No estrés	120	Delta, teta, baja alpha, alta alpha, alta beta, baja beta, gamma alta y gamma baja	Promedio de amplitud EEG predictor con mayor peso delta.ECG	Fp1 y Fp2	MIST	Corteza prefrontal

Los sujetos fueron 110 adquiridos por base de datos públicas homogeneizadas [24][25][26][27] en datos y preprocesamiento de 20 personas para su adquisición de ECG y EEG. Generando de esta forma un sistema de monitoreo clasificatorio y predictivo de nivel de estrés. Esto implica la recopilación de datos, almacenamiento, interpretación de datos y, finalmente, la validación de datos. Con el fin de generar un análisis del conjunto de datos apoyado por modelos estadísticos, de aprendizaje automatizado y técnicas de minería de datos. Obteniendo un sistema que clasifique señales EEG y ECG de acuerdo con los factores de frecuencia e intervalo de tiempo, para proceder una categorización apoyado de técnicas estadísticas. Con el fin de modelar matemáticamente la tendencia que tiene una persona a ser afectada por un estímulo que le provoque estrés. Permitiendo acelerar la toma de decisiones y poder brindar una herramienta que permite servir de indicador o parámetro inicial para el personal médico en caso de considerarlo óptimo. Y esto se visualiza en la correlación y el análisis factorial, el cual se determinó dos que teta explica a las demás ondas matemáticamente y que tienen correlación, pero no una muy estrecha.

Variables resultantes del análisis para la clasificación de ECG: HR Median RR, LF, Median REL RR, sampen y pNN50.

El resultado fue el esperado porque como se explicó en el estado de arte, recomiendan separar las señales ECG y EEG. Debido a que el agregar otra disminuye el factor de eficiencia como se ve reflejado en la tabla que se presenta, la cual presenta el porcentaje de efectividad a la que se encuentra clasificada la base de datos. Recordando que de entrenamiento se tenía más de 300, 000 datos de ECG y EEG para el entrenamiento de esta y posteriormente el entrenamiento se consideraron más de 1200 datos para clasificar el resultado de implementar este algoritmo de inteligencia se ve reflejado en los porcentajes. Añadiendo que se realizó la prueba del algoritmo como señal individual y posteriormente en señal en conjunto y como se anticipaba esta disminuye la efectividad. Al final se generó los siguientes reportes en Power BI, los cuales se presentan a continuación:

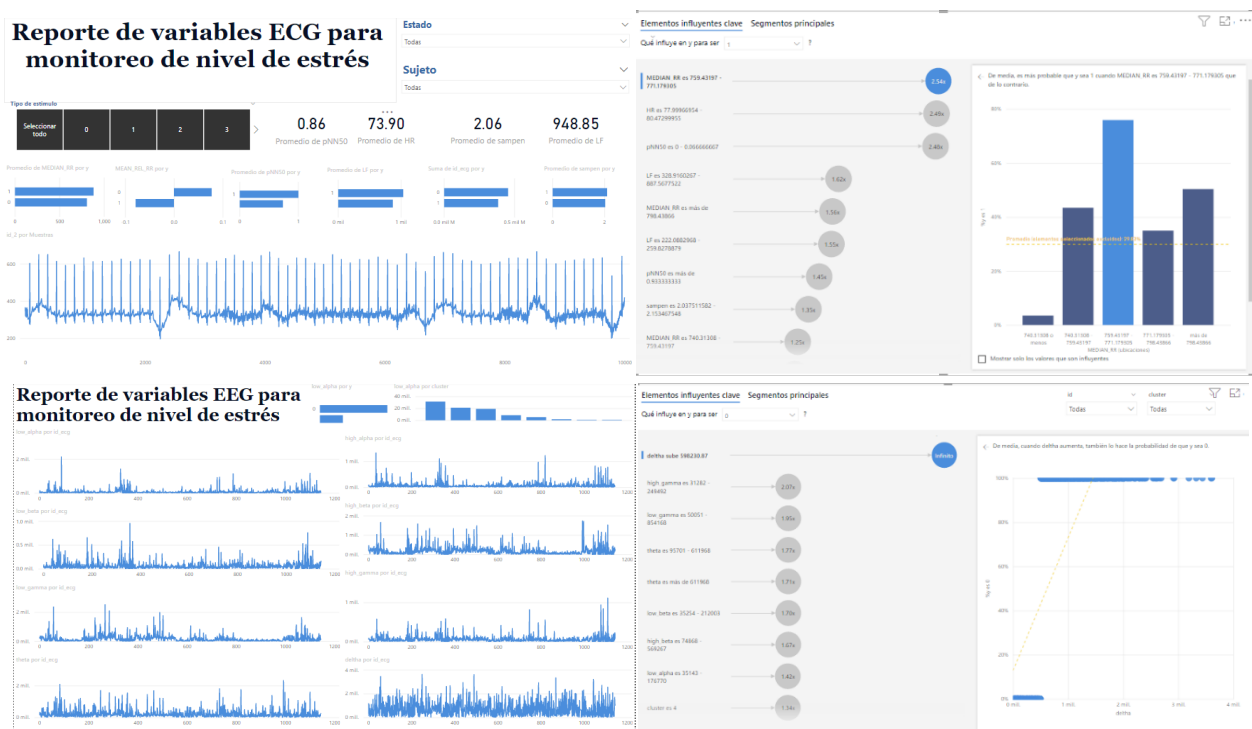


Figura 20 Presentación Power BI reporte global

La estimación de la consistencia interna se llevó a cabo con el cálculo de los coeficientes de alfa de Cronbach y de omega de McDonald. La omega de McDonald estima con mayor precisión consistencia interna n que el coeficiente alfa de Cronbach en los casos que se incumple el principio de tal equivalencia; lo anterior quiere decir que los ítems presentan coeficientes disímiles en una matriz de coeficientes de análisis factorial. De esa forma se incluyó en el modelo se incluyó la prueba de EPP 10 y se desarrolló una categorización de 8 niveles estrés. Tomando como base sus variables fisiológicas de EEG y ECG. Durante un promedio de 20 segundos cada muestra, lo cual ayudo para eliminar factores de acondicionamiento como lo es el muestreo de la señal o filtrarla.

7.1 Análisis de Resultados

Se destaca de esta recopilación los trabajos en donde la onda delta como el predictor principal de la clasificación. Ya que esta tabla tuvo gran importancia en el trabajo permitiendo esta retroalimentación eligiendo el método adecuado debido a la variabilidad que se tiene al implementar los diferentes dispositivos EEG y ECG. Lo que se consiguió en este trabajo fue El algoritmo que se presenta en este trabajo se buscará que descubra las características neurológicas relevantes de los patrones de comportamiento para la resolución de problemas. Pero se cambiará el enfoque para que el aprendizaje y la discriminación que realice se enfoque en las señales que previamente se consideren como un estímulo que genere estrés. Se buscará implementar un algoritmo en que no sea necesario ningún conocimiento de expertos neurofisiológicos para discriminar entre dos patrones de comportamiento para resolver problemas.

Se ejecutará la misma caracterización de dos patrones de comportamiento de resolución de problemas con la ayuda de arquitecturas de aprendizaje profundo. Usando un sensor electroencefalográfico no invasivo que capturen las señales de actividad cerebral con la diferencia que no se usará un sensor suave de aprendizaje profundo como se hace en este trabajo. Delimitando lo diferencial del proyecto presentado es que se añadirá la parte de minería de datos, el cual juega un papel imperativo en el dominio médico que facilita analizar los datos masivos. Para aplicar las mejores prácticas para futuras investigaciones y evaluar los informes de los pacientes y es por eso por lo que se utilizará SQL server y técnicas de minería de datos para su clasificación y predicción. Garantizando de esa manera un mejor almacenamiento de datos que se le pueda brindar al área médica con una técnica de clasificación haciendo empleo de una red neuronal multicapa.

7.2 Interpretación de Resultados

La primera interpretación de nuestros datos existe en el caso de la variable delta porque es la que explica el comportamiento de nuestros datos, según el análisis factorial y PCA. Efectos como el sueño deficiente, problemas para activar y revitalizar el cuerpo y la mente están relacionadas con la onda Teta. Por otro lado, las ondas de Teta en nuestro modelo matemático indica que al mostrar una elevada actividad cuando experimentamos emociones muy profundas. Un pico elevado de ondas Teta puede relacionarse con algún trastorno depresivo, falta de atención, mientras que los picos bajos cursan con ansiedad, estrés y baja autoconciencia emocional. Encontrando un problema debido a la baja correlación que no permite explicar de forma directa el comportamiento de cada onda EEG.

Un nivel adecuado de ondas Teta favorece la creatividad, la conexión emocional e incluso nuestra intuición. Por lo que en primera instancia se puede mencionar que el efecto del sesgo del tamaño de la muestra en la estimación de la conectividad. Este problema se puede resolver igualando los tamaños de muestra entre diferentes condiciones o participantes, utilizando estadísticas, métodos que explican explícitamente el sesgo del tamaño de la muestra, o que emplean enfoques de conectividad que no se ven afectados por el sesgo del tamaño de la muestra. Además, comprender las relaciones entre el estrés mental y las complejas y diversas características del EEG, como el tiempo, conexiones cerebrales variables, funcionales y dinámicas, necesita la integración de varios métodos de análisis de datos presentados en la descripción de estos. De esta manera se permite entenderlos de una mejor manera. Donde, como se estableció en primera instancia, se adelantó en el estado de arte combinar la conectividad de red medidas con aprendizaje para aumentar la precisión de la evaluación de los niveles de estrés mental, aunque no se alcanzan eficacias mayores al 95 % como en otros trabajos.

En conclusión, la categorización de estrés fue ocho, este no fue una categorización subjetiva o propuesta, resultado de aplicar el método mencionado en el desarrollo para conocer la agrupación por elementos con características cercanas entre ellos utilizando componentes principales.

Tabla 12 Características clúster analizando existencia de estrés en las señales ECG y EEG

Clúster				
Clúster	No estrés	Estrés	No estrés	Estrés
Uno	25522	787	50	1
Dos	7278	5370	256	182
Tres	4764	1314	603	81
Cuatro	1388	11116	688	321
Cinco	22825	20875	305	790
Seis	3762	7578	0	24
Siete	8554	17710	84	252

La anotación con el número cero es señal sin estrés y 1 es una señal con características físicas por la que pueden ser considerada estrés. Además de generar un reporte personalizado para cada persona que sus datos fueron adquiridos generados de la siguiente manera:



Figura 21 Reporte de elementos personalizados.

Clasificado de estado:

- Clúster uno definido muy alejado de características de estrés (muy relajado).
- Clúster dos definidos por estado todo bajo de control.
- Clúster tres definidos como estado neutral.
- Clúster cuatro pres-estrés (tiene características para ser delimitado como estrés en ECG y EEG).
- Clúster quinto afectado inesperadamente.
- Clúster sexto definido estresado o nervioso.
- Clúster sexto dificultades se acumulan (alto nivel de estrés).

Capítulo 7 Discusión

Las características diferenciales de este proyecto respecto a otros referenciados en el estado de arte, no solo clasifica si el sujeto tiene estrés. Añade como propuesta el poder categorizar de manera no subjetiva. Al realizar esta clasificación surgen desafíos en el ámbito psicológico, debido a que existen protocolos con estándares éticos establecidos para la experimentación con humanos respetando los tratados de Helsinki. Pueden ser indicadores del estrés percibido por el sujeto, no estiman el nivel de cortisol en estos ensayos psicológicos. A manera de valorar este factor en una metodología única que considere esta variable, se eligió el protocolo MIST, el cual permite la inducción de estrés. Generando una manipulación de activación emocional, como variable independiente, observar su efecto sobre la conducta; no obstante, esta inducción debe ser suficiente, generar incrementos significativos en los niveles de cortisol y cumplir los estándares éticos establecidos para la experimentación con humanos respetando los tratados de Helsinki.

Se solicitó el apoyo de un profesional esta área que evaluó la aplicación del proceso de registro de señal y el resultado de la etiqueta del algoritmo en un paciente. Encontró el sistema congruente con lo que él detectaba. Añadiendo que este trabajo aplicaba para las respuestas fisiológicas emitidas ante situaciones que producen un estímulo estresante. Está relacionada con la actividad del sistema nervioso autónomo (SNA) y del eje hipotálamo-hipófisis-adrenal (HPA). Cuando se hace mención del estrés agudo como del crónico, el efecto previo mediado por la rama simpática (noradrenérgica) del SNA. Esta activación inicial es continua, posterior a un corto período de tiempo del eje HPA. Si es pasajera, debido que acaba al resolverse la demanda ambiental, esto de le denomina fenómeno agudo. Considerando este factor, la base de datos pública etiqueta que las muestras presentaban un estímulo y sus características de señal (aunque no las condiciones iniciales del individuo, lo que ocasionó un sesgo en nuestro modelo).

En consecuencia, este trabajo parte de ese supuesto al generar características similares en las señales para categorizarlas de forma general. Mencionando que dos momentos del estrés median el denominado síndrome de adaptación. A causa de esto, el primer planteamiento de generar una etiqueta con un algoritmo supervisado carecía de sentido debido a que solo se cuenta con estos dos momentos y categorizarlos sin métodos no supervisados sería una evaluación superficial y subjetiva. Sin embargo, este trabajo no está capacitado para determinar cuando el estímulo estresante es persistente, debido a la activación del HPA es mantenida en el tiempo, en cuyo caso se considerará un fenómeno crónico. Algo que el algoritmo de momento está desarrollado con el fin de detectar o categorizar, sin embargo, no tiene un análisis histórico. Ese sería un área de oportunidad de este trabajo, al igual que adaptarlo a aplicaciones en donde se necesite cuantificar el avance midiendo el estrés en un tratamiento clínico o terapia.

Capítulo 8 Conclusiones

Se establece un procedimiento de aprendizaje automático para un sistema de adquisición de la muestra de pocos canales efectivo. En los elementos establecidos se estima como buena medida de preprocesamiento discriminar los datos por la distancia que se encuentra. Así es válido el poder establecer el primer parámetro descriptivo, un árbol de decisión que nos permitió identificar los predictores más importantes. Debido a la efectividad de este, como se muestra en la tabla correspondiente, se acercó a estos. Todo esto se logró por el análisis de los datos y el preprocesamiento. Las condiciones se establecieron para sistema de monitoreo en Python capaz de guardar los elementos de forma relacional. Ayudó a brindar una mejor presentación de los factores al usuario final y una manera practica de analizar para el personal médico. Al igual se logró corroborar que combinar dos variables fisiológicas como la señal ECG y EEG, sesga el modelo que lo disminuye la efectividad. Dependiendo los factores de adquisición y el tipo de experimento para determinar el nivel de estrés puede encaminar, ha resultado contradictorios con otros trabajos realizados en cuanto al peso de las variables con los predictores.

Lo anterior expuesto, hace evidente que existe un problema de unificación del criterio para cuantificar los parámetros o variables que provoquen un estímulo de estrés en una persona. Por lo que el siguiente paso sería establecer una base de datos centrada en un solo usuario, determinar sus parámetros necesarios, delimitar qué está pasando por un estímulo estresante y adecuar el aprendizaje. Comparándolo con el que combina el resultado de 110 personas, por lo que delimitar los predictores de estrés general fue un reto que se arregló con el árbol de decisión al observar su comportamiento a diferencia de las técnicas de aprendizaje supervisado. Este trabajo se diferencia en hacer un análisis predictivo del nivel de estrés, en la búsqueda de complementar software y la estructura SQL, para generar un sistema robusto de categorización multi nivel. Buscando encontrar un factor determinante al delimitar el uso y su aplicación que averigua la versatilidad en el análisis descriptivo para presentar un reporte global.

Utilizar un sistema matemático analítico para poder delimitar el accionar predictivo de estas variables con el fin de establecer los mejores predictores fue fundamental para definir los parámetros de nuestro sistema de adquisición. Aunque al juntar las dos señales fisiológicas, el resultado final de la predicción fuera desfavorable debido a que no mantienen una correlación directa en ellas. El estrés se entiende desde tres perspectivas: la primera es la ambiental, que se centra en los eventos vitales estresores; la segunda, la psicológica, que implica la experiencia subjetiva y emocional frente a los estímulos; y la tercera, la biomédica, que estudia la respuesta y los sistemas fisiológicos involucrados en el afrontamiento de los asuntos vitales. Debido a los tres factores se desarrolló un modelo de aprendizaje no supervisado que implique categorizar y evitar subjetividad. De una forma se agruparon registros por sus características y se interpretó la representación. Con el fin de corregir o modificar el programa, este se añadió en la parte del anexo C el código que sirva como base a futuros proyectos a la comunidad.

Capítulo 10 Recomendaciones

La base de datos se obtuvo con 20 sujetos que cumplieran con todos los criterios que se mencionaron en el protocolo de pruebas y que por decisión propia quisieron participar. La muestra no es representativa a la población de estudio, sin embargo, se considera un buen comienzo el desarrollo de este algoritmo de categorización del nivel de estrés. Puede decirse que el protocolo diseñado e implementado permitió realizar una prueba organizada y metódica, de tal manera que se cumplió el objetivo de generar un estresante en los momentos deseados. Recopilando bases de datos públicas con la desventaja de ausencia de condiciones en que se encontraba la persona que suponen un sesgo en el modelo, como la edad, test, personalidad, estrés percibido, condición física, etc. Elaborar la base de datos con señales propias de este proyecto requirió ajustar los parámetros que se tenían para generar una unificación de estas.

Sobre la base de datos de señales ECG debe de mencionarse que no tiene un fin médico el módulo, por lo que su calidad no es la misma y presenta ruido y alteraciones, con la que no fue posible reducirlo con el procesamiento y filtrado. Aunque se buscó registrar y acondicionar mejor señal para adquirir las seis características necesarias, en su análisis. El EEG, la validez funcional de electroencefalografía (EEG) de bajo costo, aún está en debate. En comparación en profundidad de los registros de EEG que han obtenido con un dispositivo ambulatorio de electrodos de copa dorada de grado médico, el SOMNOwatch + EEG-6, frente a los de un electrodo seco único de bajo costo de alcance de consumo. Presenta una concordancia sustancial entre las señales: mientras que, NeuroSky MindWave muestra niveles más altos de ruido y una forma bifásica de parpadeos, la métrica de similitud indicó que las señales de ambos estaban significativamente correlacionadas. Si bien NeuroSky MindWave era menos confiable, los dispositivos tenían una tasa de detección de parpadeo similar. En general, tiene un ruido limitado, pero proporciona grabaciones estables incluso durante largos períodos de tiempo. Además, sus datos serían de calidad adecuada en comparación con los dispositivos EEG de electrodo húmedo convencionales, excepto por un posible error de calibración y diferencias espectrales a bajas frecuencias. Presenta una concordancia sustancial entre las señales: mientras que, NeuroSky MindWave era menos confiable, ambos dispositivos tenían una tasa de detección semejante. Presentó niveles más altos de ruido y una manera bifásica de parpadeos, la métrica de similitud indicó que las señales de ambos dispositivos de grabación tenían una tasa de detección semejante [34]. Es importante considerar esta parte si en un futuro se trabajara con la base de datos o el dispositivo de adquisición por el sesgo que puede presentar los datos al mejorar el dispositivo de adquisición.

Bibliografía

- [1] HHRI [Internet]. Nota informativa provisional: Cómo abordar la salud mental y los aspectos psicosociales del brote de COVID-19 - HHRI; [consultado el 6 de diciembre de 2021]. Disponible en: <https://www.hhri.org/es/publication/nota-informativa-provisional-como-abordar-la-salud-mental-y-los-aspectos-psicosociales-del-brote-de-covid-19/>.
- [2] C L YH. Escalas de Apreciación del Estrés y sus Influencia en la Sociabilidad del personal a cargo del Banco de Sangre en la Cruz Roja Ecuatoriana Sede Ambato. Ecuador: Bachelor's thesis, Universidad Técnica de Ambato-Facultad de Ciencias de la Salud-Carrera Psicología Clínica; 2016.
- [3] Psicología y Salud [Internet]. Estrés académico en estudiantes universitarios | Psicología y Salud; [consultado el 6 de diciembre de 2021]. Disponible en: <https://psicologiaysalud.uv.mx/index.php/psicysalud/article/view/584>
- [4] X G. Faster Detection of Abnormal Electrocardiogram (ECG) Signals Using Fewer Features of Heart Rate Variability (HRV). Journal of Computer Science & Systems Biology [Internet]. 2018 [consultado el 6 de diciembre de 2021];12(01). Disponible en: <https://doi.org/10.4172/jcsb.1000295>
- [5] Amira T, Dan I, Atta B, Said G, Az-eddine B, Katarzyna WW. Stress Level Classification Using Heart Rate Variability. Advances in Science, Technology and Engineering Systems Journal [Internet]. 2019 [consultado el 6 de diciembre de 2021];4(3):38-46. Disponible en: <https://doi.org/10.25046/aj040306>
- [6] Saucedo-Villanueva EG, Flores-Sánchez AC, Villanueva-Cruz LD, García-Sánchez AA, Cardona-Perales DH, Martín del Campo-Cervantes J. Efectos de la actividad física moderada sobre la ingesta alimentaria en presencia de estrés psicológico. Lux Médica [Internet]. 31 de enero de 2016 [consultado el 6 de diciembre de 2021];11(32):09-14. Disponible en: <https://doi.org/10.33064/32lm2016731>
- [7] Upadhyay PK, Sinha RK, Karan BM. Detection and analysis of the effects of heat stress on EEG using wavelet transform —EEG analysis under heat stress. Journal of Biomedical Science and Engineering [Internet]. 2010 [consultado el 6 de diciembre de 2021];03(04):405-14. Disponible en: <https://doi.org/10.4236/jbise.2010.34056>
- [8] O'Connor DB, Thayer JF, Vedhara K. Stress and Health: A Review of Psychobiological Processes. Annual Review of Psychology [Internet]. 4 de enero de 2021 [consultado el 6 de diciembre de 2021];72(1):663-88. Disponible en: <https://doi.org/10.1146/annurev-psych-062520-122331>
- [9] Shah NH, Milstein A, Bagley, PhD SC. Making Machine Learning Models Clinically Useful. JAMA [Internet]. 8 de octubre de 2019 [consultado el 6 de diciembre de 2021];322(14):1351. Disponible en: <https://doi.org/10.1001/jama.2019.10306>

- [10] Zainudin Z, Hasan S, Shamsuddin SM, Argawal S. Stress Detection using Machine Learning and Deep Learning. Journal of Physics: Conference Series [Internet]. 1 de agosto de 2021 [consultado el 6 de diciembre de 2021];1997(1):012019. Disponible en: <https://doi.org/10.1088/1742-6596/1997/1/012019>
- [11] Encyclopedia of medical devices & instrumentation. 2a ed. Hoboken, NJ: Wiley; 2006.
- [12] Electrocardiografía clínica. De lo básico a lo complejo. [lugar desconocido]: Distribuna; 2014.
- [13] Electrocardiografía: Interpretación práctica del ECG. [lugar desconocido]: Editorial Médica Panamericana; 2015.
- [14] Tichon JG, Wallis G, Riek S, Mavin T. Physiological measurement of anxiety to evaluate performance in simulation training. Cognition, Technology & Work [Internet]. 9 de marzo de 2013 [consultado el 6 de diciembre de 2021];16(2):203-10. Disponible en: <https://doi.org/10.1007/s10111-013-0257-8>
- [15] Physiological Stress Prediction using Machine Learning Classifiers. International Journal of Innovative Technology and Exploring Engineering [Internet]. 10 de noviembre de 2019 [consultado el 6 de diciembre de 2021];9(1):675-7. Disponible en: <https://doi.org/10.35940/ijitee.a4556.119119>
- [16] Disha Sharma et al, DS. Stress Prediction of Students using Machine Learning. International Journal of Mechanical and Production Engineering Research and Development [Internet]. 2020 [consultado el 6 de diciembre de 2021];10(3):5609-20. Disponible en: <https://doi.org/10.24247/ijmperdjun2020534>
- [17] Núñez E, Steyerberg EW, Núñez J. Estrategias para la elaboración de modelos estadísticos de regresión. Revista Española de Cardiología [Internet]. Junio de 2011 [consultado el 6 de diciembre de 2021];64(6):501-7. Disponible en: <https://doi.org/10.1016/j.recesp.2011.01.019>
- [18] Barragán A, Ruiz CH, Gómez Vargas E. DISEÑO DE UNA APLICACIÓN ADAPTATIVA PARA MONITOREO REMOTO A TRAVÉS DE TECNOLOGÍAS MÓVILES. Redes de Ingeniería [Internet]. 22 de mayo de 2012 [consultado el 6 de diciembre de 2021];1(1):43. Disponible en: <https://doi.org/10.14483/2248762x.7149>
- [19] Jebelli H, Hwang S, Lee S. EEG-based workers' stress recognition at construction sites. Automation in Construction [Internet]. Septiembre de 2018 [consultado el 6 de diciembre de 2021];93:315-24. Disponible en: <https://doi.org/10.1016/j.autcon.2018.05.027>
- [20] Schroder P. Adversity, Stress, and Psychopathology. Journal of Psychosocial Nursing and Mental Health Services [Internet]. Abril de 1999 [consultado el 6 de diciembre de 2021];37(4):49. Disponible en: <https://doi.org/10.3928/0279-3695-19990401-20>.
- [21] Villalba-Diez J, Zheng X, Schmidt D, Molina M. Characterization of Industry 4.0 Lean Management Problem-Solving Behavioral Patterns Using EEG Sensors and Deep Learning.

Sensors [Internet]. 26 de junio de 2019 [consultado el 6 de diciembre de 2021];19(13):2841. Disponible en: <https://doi.org/10.3390/s19132841>

[22] Flesia L, Monaro M, Mazza C, Fietta V, Colicino E, Segatto B, Roma P. Predicting Perceived Stress Related to the Covid-19 Outbreak through Stable Psychological Traits and Machine Learning Models. Journal of Clinical Medicine [Internet]. 19 de octubre de 2020 [consultado el 6 de diciembre de 2021];9(10):3350. Disponible en: <https://doi.org/10.3390/jcm9103350>

[23] S T. Stress Analysis using EEG signals. International Journal for Research in Applied Science and Engineering Technology [Internet]. 31 de mayo de 2019 [consultado el 6 de diciembre de 2021];7(5):1281-5. Disponible en: <https://doi.org/10.22214/ijraset.2019.5213>

[24] Kaggle: Your Machine Learning and Data Science Community [Internet]. Biometrics for stress monitoring; [consultado el 6 de diciembre de 2021]. Disponible en: <https://www.kaggle.com/qiri/stress>.

[25] PhysioNet [Internet]. PhysioNet Databases; [consultado el 6 de diciembre de 2021]. Disponible en: <https://physionet.org/about/database/>.

[26] PhysioNet [Internet]. MIT-BIH Noise Stress Test Database; 3 de agosto de 1999 [consultado el 6 de diciembre de 2021]. Disponible en: <https://physionet.org/content/nsttdb/1.0.0/>.

[27] MIT Media Lab [Internet]. Data – MIT Media Lab; [consultado el 6 de diciembre de 2021]. Disponible en: <https://www.media.mit.edu/groups/affective-computing/data/>.

[28] Khan Academy [Internet]. Khan Academy | Práctica, lecciones y cursos en línea gratuitos; [consultado el 15 de diciembre de 2021]. Disponible en: <https://es.khanacademy.org/>.

[29] ResearchGate [Internet]. ResearchGate | Find and share research; [consultado el 15 de diciembre de 2021]. Disponible en: <https://www.researchgate.net>

[30] Enfermería [Internet]. Enfermería; [consultado el 15 de diciembre de 2021]. Disponible en: <https://enfermeria.top>

[31] UNIT Electronics [Internet]. Tarjetas de Desarrollo - UNIT Electronics; [consultado el 15 de diciembre de 2021]. Disponible en: <https://blog.uelectronics.com/tarjetas-desarrollo>

[32] Amazon.com [Internet]. Amazon.com; [consultado el 15 de diciembre de 2021]. Disponible en: <https://www.amazon.com>

[33] Createc 3D - Posiblemente los mejores precios en impresion 3D y robotica - CREATEC4 S.C.A. [Internet]. Createc 3D - Posiblemente los mejores precios en impresion 3D y robotica - CREATEC4 S.C.A.; [consultado el 15 de diciembre de 2021]. Disponible en: <https://createc3d.com>

[34] Rieiro H, Diaz-Piedra C, Morales JM, Catena A, Romero S, Roca-Gonzalez J, et al. Validation of electroencephalographic recordings obtained with a consumer-grade, single dry electrode, low-cost device: A comparative study. *Sensors (Basel)* [Internet]. 2019 [cited 2022 May 24];19(12):2808. Available from: <http://dx.doi.org/10.3390/s19122808>

[35] Campo-Arias A, Pedrozo-Cortés MJ, Pedrozo-Pupo JC. Escala de estrés percibido relacionado con la pandemia de COVID-19: una exploración del desempeño psicométrico en línea. *Rev Colomb Psiquiatr (Engl)* [Internet]. 2020 [cited 2022 May 24];49(4):229–30. Available from: <http://dx.doi.org/10.1016/j.rcp.2020.05.005>

[36] Campo-Arias A, Oviedo HC, Herazo E. Escala de Estrés Percibido-10: Desempeño psicométrico en estudiantes de medicina de Bucaramanga, Colombia. *Rev Fac Med Univ Nac Colomb* [Internet]. 2015;62(3):407–13. Available from: <http://www.scielo.org.co/pdf/rfmun/v62n3/v62n3a10.pdf>

Referencia de código:

[37] Autoencoders: explicación y tutorial en Python [Internet]. Codificando Bits. [cited 2022 May 27]. Available from: <https://www.codificandobits.com/blog/autoencoders-explicacion-y-tutorial-python/>

[38] Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4), 1689–1696. <https://doi.org/10.3758/s13428-020-01516-y>

[39] Arboles de decision python [Internet]. Cienciadedatos.net. [cited 2022 May 27]. Available from: https://www.cienciadedatos.net/documentos/py07_arboles_decision_python.html

Anexo A Protocolo de adquisición de señales

Para el protocolo de adquisición de señales se llevó a cabo en las 10 personas que se recolectaron la base de datos. Primero comenzó el reconocimiento de este con la siguiente carta de consentimiento.

CONSENTIMIENTO INFORMADO PARA EL REGISTRO DE SEÑALES EEG Y ECG MONITOREO DE ESTRÉS

Yo, C. _____ en pleno uso de mis facultades mentales, sin ninguna presión o inducción de cualquier tipo, informo que he leído, comprendido y aclarado todas mis dudas respecto a la información del estudio.

Acepto participar de manera voluntaria en los procedimientos llevados a cabo, en el entendimiento de que no afectarán negativamente mi estado de salud.

Asumo la responsabilidad de responder con veracidad a todos los cuestionarios realizados a mi persona y cumplir con las condiciones descritas para el estudio, informando el incumplimiento a alguna de ellas en el entendimiento de que estas pueden afectar al estudio.

Doy mi consentimiento para ser contactado periódicamente aún después de la terminación de la primera fase con el fin de proveer información de utilidad a los procedimientos realizados y entiendo que puedo terminar mi participación en cualquier momento con la apropiada notificación.

Ciudad y Fecha: _____

Firma del participante

Yo, **Ian Felipe Jaime Hernandez**, he explicado los objetivos del estudio y los procedimientos que se llevarán a cabo al C. _____ con énfasis en los riesgos y beneficios de su participación, aclarando todas sus dudas.

Ciudad y Fecha: _____

Firma del investigador

Jurado del proyecto terminal UPIITA

Del Proyecto con registro TTB-2020/1-23

De: Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas UPIITA - IPN

Asunto: Formalización de participación en la retroalimentación del Proyecto TTB-2020/1-23

Cortésmente, deseamos formalizar mi participación activamente en el desarrollo del Proyecto de Sistema de monitoreo no invasivo del nivel de estrés a partir de señales EEG y ECG.

Participa en reflexiones, propuestas y/o acciones desde mi área de enfoque profesional que promovieron la cooperación con respecto, la solidaridad y la cooperación mutua para realizar retroalimentación y comentarios al trabajo realizado. Trabajamos específicamente sobre los temas de:

Escala en la medición de estrés desde el punto de vista de la terapia, psicología y psicométrico.

Luego de haber leído el documento de proyecto estamos de acuerdo con las conclusiones llegadas.

Atentamente,



Mtro. Josafat Jonathan Sánchez Dueñas

Firma y Nombre

Maestro en Psicología con residencia en Neuropsicología Clínica

Cedula Profesional: 08708703 / 10158715

En época de coronavirus lo que se llevó a cabo fue las medidas estándares en el proceso de adquisición de datos:

Higiene de manos y protocolo para la higiene respiratoria

Se recomendó que el voluntario a adquirir la señal y los asistentes se laven las manos con frecuencia (por ejemplo, antes, durante y después de comenzar el protocolo, o después de tocar los instrumentos) con agua y jabón por al menos 20 segundos, y aumente el monitoreo para garantizar el cumplimiento de estas medidas.

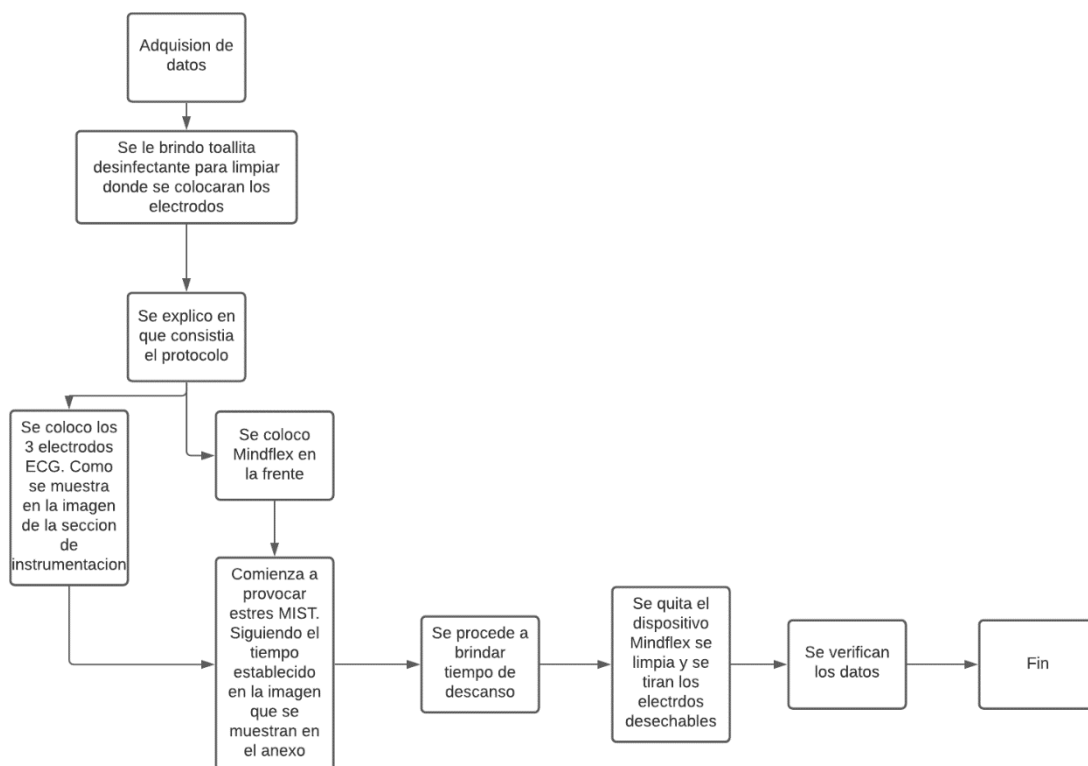
Uso de desinfectante de manos antes y después de adquisición de datos que contenga al menos un 60 % de alcohol y frotarse las manos hasta que se sequen.

Se les recomendó a los invitados que eviten cantar o gritar, especialmente en espacios cerrados como el que se adquieren los datos.

Suministros adecuados y uso de mascarilla en todo momento.

Se aseguro de contar con los suministros adecuados para alentar hábitos de higiene. Los suministros incluyen jabón, agua, desinfectante de manos con al menos un 60 % de alcohol, un método para secarse las manos (p. ej., toallas de papel, secadores de aire para manos), pañuelos desechables, toallitas desinfectantes, mascarillas (según sea posible) y botes de basura que no requieran contacto con las manos.

Posteriormente después de establecer los dispositivos:



En cuanto a tiempo de descanso se siguió el implementado en la base de datos previa:

	B	C	D	E	F	G	H	I	J	K	L
1	PP	C	timestamp	HR	RMSSD	SCL	date	subject	label	Condition	ElapsedTime
2	PP18		120121019T092600000	81	0.038677	399.218169	2012-10-19 09:26:00	p18	rest	R	0
3	PP18		120121019T092700000	83	0.061672	442.392008	2012-10-19 09:27:00	p18	rest	R	1
4	PP18		120121019T092800000			488.771771	2012-10-19 09:28:00	p18	rest	R	2
5	PP18		120121019T092900000	88	0.018767	617.065404	2012-10-19 09:29:00	p18	rest	R	3
6	PP18		120121019T093000000	87	0.027548	501.635419	2012-10-19 09:30:00	p18	rest	R	4
7	PP18		120121019T093100000	89	0.022025	518.255352	2012-10-19 09:31:00	p18	rest	R	5
8	PP18		120121019T093700000	73	0.035065	235.700016	2012-10-19 09:37:00	p18	no stress	N	11
9	PP18		120121019T093800000	84	0.028728	227.291769	2012-10-19 09:38:00	p18	no stress	N	12
10	PP18		120121019T093900000			224.73756	2012-10-19 09:39:00	p18	no stress	N	13
11	PP18		120121019T094000000			236.44336	2012-10-19 09:40:00	p18	no stress	N	14
12	PP18		120121019T094100000			233.905076	2012-10-19 09:41:00	p18	no stress	N	15
13	PP18		120121019T094200000	86	0.022428	251.544692	2012-10-19 09:42:00	p18	no stress	N	16
14	PP18		120121019T094300000	82	0.019159	287.931209	2012-10-19 09:43:00	p18	no stress	N	17
15	PP18		120121019T094400000	83	0.023304	294.442903	2012-10-19 09:44:00	p18	no stress	N	18
16	PP18		120121019T094500000			249.629473	2012-10-19 09:45:00	p18	no stress	N	19
17	PP18		120121019T094600000	78	0.17013	250.957704	2012-10-19 09:46:00	p18	no stress	N	20
18	PP18		120121019T094700000	84	0.026012	242.107298	2012-10-19 09:47:00	p18	no stress	N	21
19	PP18		120121019T094800000	90	0.047499	210.597075	2012-10-19 09:48:00	p18	no stress	N	22
20	PP18		120121019T094900000			210.349003	2012-10-19 09:49:00	p18	no stress	N	23
21	PP18		120121019T095000000	81	0.028766	214.239902	2012-10-19 09:50:00	p18	no stress	N	24
22	PP18		120121019T095100000	80	0.026084	218.972546	2012-10-19 09:51:00	p18	no stress	N	25
23	PP18		120121019T095200000	81	0.031017	216.694338	2012-10-19 09:52:00	p18	no stress	N	26
24	PP18		120121019T095300000	91	0.029545	206.359091	2012-10-19 09:53:00	p18	no stress	N	27
25	PP18		120121019T095400000	74	0.037645	241.531843	2012-10-19 09:54:00	p18	no stress	N	28
	B	C	D	E	F	G	H	I	J	K	L
29	PP18		120121019T095800000			230.281217	2012-10-19 09:58:00	p18	no stress	N	32
30	PP18		120121019T095900000	80	0.04912	219.134303	2012-10-19 09:59:00	p18	no stress	N	33
31	PP18		120121019T100000000	78	0.032602	215.070435	2012-10-19 10:00:00	p18	no stress	N	34
32	PP18		120121019T100100000			230.691926	2012-10-19 10:01:00	p18	no stress	N	35
33	PP18		120121019T100200000			209.015149	2012-10-19 10:02:00	p18	no stress	N	36
34	PP18		120121019T100300000	76	0.040157	208.460842	2012-10-19 10:03:00	p18	no stress	N	37
35	PP18		120121019T100400000			204.84877	2012-10-19 10:04:00	p18	no stress	N	38
36	PP18		120121019T100500000	75	0.04055	206.637342	2012-10-19 10:05:00	p18	no stress	N	39
37	PP18		120121019T100600000	85	0.029004	202.795393	2012-10-19 10:06:00	p18	no stress	N	40
38	PP18		120121019T100700000	81	0.103429	209.528614	2012-10-19 10:07:00	p18	no stress	N	41
39	PP18		120121019T100800000			206.758538	2012-10-19 10:08:00	p18	no stress	N	42
40	PP18		220121019T102600000			227.734168	2012-10-19 10:26:00	p18	rest	R	60
41	PP18		220121019T102700000	70	0.052653	262.902151	2012-10-19 10:27:00	p18	rest	R	61
42	PP18		220121019T102800000	73	0.037109	296.553511	2012-10-19 10:28:00	p18	rest	R	62
43	PP18		220121019T102900000	77	0.041778	347.582743	2012-10-19 10:29:00	p18	rest	R	63
44	PP18		220121019T103000000	79	0.031774	346.602554	2012-10-19 10:30:00	p18	rest	R	64
45	PP18		220121019T103100000	71	0.039093	400.666286	2012-10-19 10:31:00	p18	rest	R	65
46	PP18		220121019T103400000	77	0.035159	275.207487	2012-10-19 10:34:00	p18	interruptio	I	68
47	PP18		220121019T103500000	77	0.033185	262.258636	2012-10-19 10:35:00	p18	interruptio	I	69
48	PP18		220121019T103600000	77	0.024955	295.297823	2012-10-19 10:36:00	p18	interruptio	I	70
49	PP18		220121019T103700000	77	0.029211	276.393039	2012-10-19 10:37:00	p18	interruptio	I	71
50	PP18		220121019T103800000	78	0.026955	275.809922	2012-10-19 10:38:00	p18	interruptio	I	72
51	PP18		220121019T103900000	76	0.038782	283.459302	2012-10-19 10:39:00	p18	interruptio	I	73
52	PP18		220121019T104000000	75	0.037341	247.533453	2012-10-19 10:40:00	p18	interruptio	I	74
53	PP18		220121019T104100000	77	0.032674	223.952283	2012-10-19 10:41:00	p18	interruptio	I	75
54	PP18		220121019T104200000	73	0.03986	227.380228	2012-10-19 10:42:00	p18	interruptio	I	76

En cuanto al método para delimitar el proceso de estrés se desarrolló un protocolo para inducir estrés psicológico moderado en un entorno de imágenes funcionales y evaluar los efectos del estrés en la fisiología y la activación cerebral. Métodos: La Tarea de estrés por imágenes de Montreal (MIST), derivada de la Prueba de desafío mental de Trier, consiste en una serie de desafíos aritméticos mentales computarizados, junto con componentes de amenaza de evaluación social que están integrados en el programa o presentados por el investigador. Para permitir que los efectos del estrés y la aritmética mental se investiguen por separado, el MIST tiene 3 condiciones de prueba (reposo, control y experimental), que pueden presentarse en un bloque o en un diseño relacionado con eventos, para su uso con imágenes de resonancia magnética funcional. (fMRI) o tomografía por emisión de positrones (PET). En estado de reposo, los sujetos miran una pantalla de computadora estática en la que no se muestran tareas. En la condición de control, una serie de tareas aritméticas mentales se muestran en la pantalla de la computadora y los sujetos envían sus respuestas por medio de una interfaz de respuesta. En la condición experimental, la dificultad y el límite de tiempo de las tareas se manipulan para que estén más allá de la capacidad mental del individuo. Además, en esta condición, la presentación de las tareas de aritmética mental se

complementa con una visualización de información sobre el rendimiento individual y medio, así como el rendimiento esperado. Al completar cada tarea, el programa presenta una evaluación de desempeño para aumentar aún más la amenaza evaluativa social de la situación.

Base de datos generados EEG

	A	B	C	D	E	F	G	H	I	J	K
1	strength	attention	meditation	delta	theta	low_alpha	high_alpha	low_beta	high_beta	low_gamma	high_gamma
2	51	0	0	1154595	179193	13335	44484	33138	204057	262653	64201
3	51	0	0	145774	322483	139989	83856	83715	461158	193445	88907
4	51	0	0	1084302	142764	60200	30645	37110	181713	328328	71307
5	51	0	0	247533	241141	23523	39242	48259	223951	362543	56752
6	80	0	0	751386	262307	16149	13767	16745	131624	185840	58791
7	51	0	0	732536	299231	61445	19438	30322	159233	295324	58968
8	51	0	0	63342	30925	13487	11084	13345	41400	47594	11804
9	51	0	0	242132	147836	15266	13274	48861	238946	356140	51123
10	80	0	0	140984	16296	15423	5302	21806	178659	74097	18713
11	51	0	0	93754	554482	79950	37446	17863	334248	270312	24529
12	51	0	0	758975	590351	59454	24458	73366	191907	234176	29790
13	51	0	0	164831	130272	46359	17951	32852	227393	289366	30535
14	80	0	0	603628	168004	29863	55523	135779	262013	348820	56366

Base de datos ECG con todas las variables estudiadas

A		B		C		D		E		F		G		H		I		J		K	
1	MEAN_RR	MEDIAN_RR	SDRR	RMSSD	SDSD	SDRR_RMSSD	HR	pNN25	pNN50	SD1	SD2										
2	721.901896873333	727.26728	74.7223152241469	12.3612638925376	12.361069336259	6.04487663023326	84.1218683005979	4.93333333333333		0	8.74351288554139	105.3109669850									
3	843.53863386667	844.40793	58.4994290575611	19.2988800459214	19.2987953740711	3.03123439900982	71.4786420208037	21		0.2	13.6508631606532	81.59669316271									
4	958.52386753333	966.671125	132.849110322794	21.3427154202397	21.3426531829438	6.22456457423456	63.8742930210046	24.1333333333333		1.8	15.0965711816958	187.2695003615									
5	824.8386688	842.485905	117.822093533661	11.7718144462828	11.7712475922034	10.0088303354854	74.3305311634378	4.73333333333333	0.533333333333333	8.32630674592406	166.4174391861										
6	756.707932713333	747.94162	143.968457491862	13.3577475646879	13.3563883797967	10.778992524498	82.0920490804156	5.93333333333333	0.666666666666667	9.4475446036444	203.3828345906										
7	831.632294506667	836.596795	62.0690416985758	12.5766666581141	12.5765570388297	4.93525378272874	72.5626970709208	4.53333333333333	0.266666666666667	8.89593655155676	87.32693849912										
8	832.89481856	834.29	84.9992018815892	14.1478045498427	14.1472570898944	6.0079428989447	72.8169597692694	7.4		0	10.0069598581607	119.789771664									
9	719.438761046667	717.521215	57.9713581170895	9.35084841225618	9.35050696468337	6.19956270750121	83.9512851239477	1.66666666666667	0.133333333333333	6.61401339221004	81.71665405978										
10	997.711161386667	998.858135	80.2155094089391	22.1167682503908	22.1167292195229	3.62890916235113	60.5375704596281	26.8666666666667	1.93333333333333	15.6441082609002	112.3579893741										
11	751.381601573334	760.687035	88.536256060048	13.9995242338353	13.9984054616803	6.3242320830187	81.1254902080319	6.73333333333333	0.6	9.90167073751542	124.8170428702										
12	751.879508473333	759.246315	83.9331890138678	14.6319802838954	14.6317746008923	5.73628363250203	80.9498293711738	7.53333333333333	0.666666666666667	10.3496798117595	118.2473871333										
13	805.624305246667	807.941995	48.3181894814931	14.7547956023457	14.7547647463751	3.27447475307619	74.7502207275062	9.8		0	10.4366760005657	67.53051653584									
HR	pNN25	pNN50	SD1	SD2	KURT	SKREW	MEAN_REL_RR	MEDIAN_REL_RR													
84.1218683005979	4.93333333333333		0	8.74351288554139	105.310966985086	1.26295815403176	-0.703778809878342	8.09903035542476E-05	-0.00095117154872												
71.4786420208037	21		0.2	13.6508631606532	81.5966931627157	-0.445805949307516	-0.144911277914089	6.13213175076386E-05	-0.001542892264451												
63.8742930210046	24.1333333333333		1.8	15.0965711816958	187.269500361569	-0.666523080618011	-0.103724783407206	-5.53297687265202E-05	-0.001999976622838												
74.3305311634378	4.73333333333333	0.533333333333333	8.32630674592406	166.417439186169	-0.698096449456366	-0.145091938781794	0.000130468603965	0.000444737720064	0.000444737720064												
82.0920490804156	5.93333333333333	0.666666666666667	9.4475446036444	203.382834590666	1.13411050260092	0.769517204620834	0.00031047768167	-0.000169636304302	-0.000169636304302												
72.5626970709208	4.53333333333333	0.266666666666667	8.89593655155676	87.3269384991238	-0.087276895329977	-0.226072011328591	5.49226166315477E-05	0.000707475010865	0.000707475010865												
72.8169597692694	7.4		0	10.0069598581607	119.78977166441	-0.30931124003355	-0.091205326725763	-0.00013074531082	-0.000265749823458												
83.9512851239477	1.66666666666667	0.133333333333333	6.61401339221004	81.7166540597883	-0.40337320330621	-0.05707744147206	-0.000101225286076	6.74117451702213E-05	6.74117451702213E-05												
60.5375704596281	26.8666666666667	1.93333333333333	15.6441082609002	112.357989374109	-0.098171249214935	-0.116414964114693	5.2887971837486E-05	-0.001117543086188	-0.001117543086188												
1	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ											
1	HF	HF_PCT	HF_NU	TP	LF_HF	HF_LF	sampen	higucl	datasetid	condition											
2	66.6170570847812	3.92186775670829	9.76028862766321	1698.60538950692	9.24559865131179	0.108159572756072	2.09734241608002	1.24369569112013		2/no stress											
3	26.5000861840195	1.1234159198523	1.66315099604706	2358.88469405913	59.1268316813555	0.016912795283691	2.21727481367058	1.25005593748288		2/time pressure											
4	16.0249347900764	0.370207687037512	0.766415522697417	4328.6337240352	129.477524317419	0.007723348166192	2.21713597639173	1.14494257420593		2/no stress											
5	17.581470340714	0.615932177300256	3.35865151872517	2854.4490906668	28.7738540132787	0.034753773322771	2.10686324732983	1.14235498899278		2/no stress											
6	35.199053927255	0.662878942044663	6.29225305341141	5310.02747178585	14.8925585400263	0.067147629288301	1.91219110606505	1.12809846180359		2/interruption											
7	27.5170120606286	1.28046967024572	4.48071503113865	2148.97804298232	21.3178665246622	0.046909093044022	2.17795578162885	1.2025433152744		2/time pressure											
8	14.7718360563585	0.594187095698687	1.53697484045159	2486.05811423567	64.0628737492007	0.015609665028686	2.18665488337716	1.18397977392098		2/interruption											
9	34.51895979392	2.19501558821071	8.97910367593813	1572.60659854951	10.136969079439	0.098648816245116	2.19555818801543	1.19236060883153		2/no stress											
10	8.95641481732311	0.245826301624386	0.519305738598294	3643.39159779466	191.564789038779	0.005270166007635	2.19719789323594	1.22477194819241		2/interruption											
RMSSD_REL_RR	SDSD_REL_RR	SDRR_REL_RR	RMSSD_REL_RR	KURT_REL_RR	SKREW_REL_RR	VLF	VLF_PCT	LF	X	Y	Z										
0.011207502597323	0.011207500770617		1.5708002503703	1.26295815403176	-0.703778809878342	1016.07375928473	59.8181170012463	615.914573137414	36.2600152420454	90.2397113723											
0.011969554410512	0.01196953358325		1.91895321604694	-0.445805949307516	-0.144911277914089	765.518472531173	32.4525600788855	1566.86613534393	66.4240240012622	98.336849003											
0.009605921451421	0.00960591489692		2.29847313060604	-0.666523080618011	-0.103724783407206	2237.73990527794	51.6962175120675	2074.86888396718	47.933574800895	99.2335844773											
0.007051152453893	0.007051141559368		2.11603303919216	-0.698096449456366	-0.145091938781794	2330.98095662398	81.6613252709813	505.886663608747	17.7227442557185	96.6413484812											
0.01168889545007	0.011688842214862		1.68103439256459	1.13411050260092	0.769517204620834	4750.6244466934	89.4651576086045	524.203971165187	9.87196344935083	93.7077469465											
0.008242854510484	0.008242834069058		1.88101720479456	-0.087276895329977	-0.226072011328591	1534.8570406557	71.4226488105779	586.603990266	27.2968815191764	95.5192849688											

Anexo B Código y documentación GitHub

Se recopilaron los programas realizados en GitHub con la base de datos de este proyecto.

Link: <https://github.com/ianherz118/TT-estres-deteccion-ian->

Donde se adjuntan base de datos y programas realizados para la presentación de base de datos e introducción a la interfaz de usuario.

Anexo C Hojas de especificaciones de dispositivos

Mechanical Drawing

The dimensions and major components of the TGAM is shown in the mechanical drawing in [Figure 1](#). There are two mounting holes at the upper right and lower left corner. They can be used to secure the TGAM to your system housing.

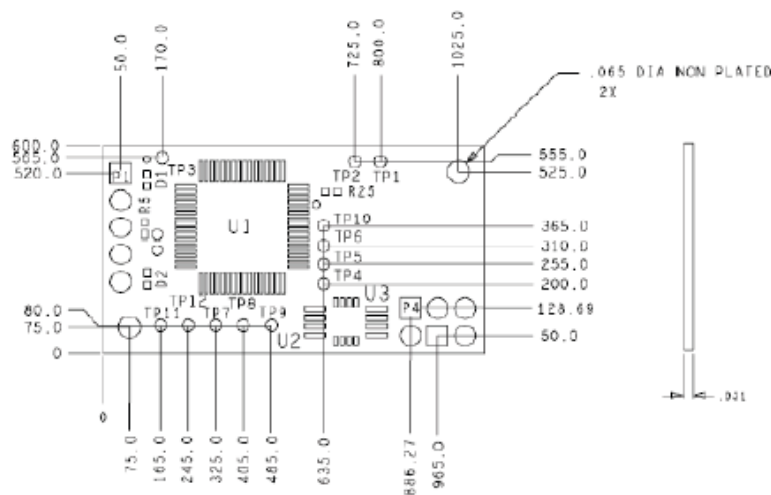


Fig 1 Mechanical Drawing & Board Thickness

Board Layout

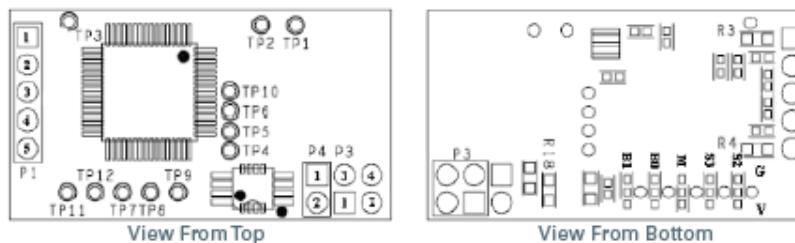


Fig 2 Board Layout Note: Labels in "" Indicated on PCB for convenience

Header p1 (Electrode)

Pin 1 - EEG Electrode "EEG"
Pin 2 - EEG Shield
Pin 3 - Ground Electrode
Pin 4 - Reference Shield
Pin 5 - Reference Electrode "REF"

Header p4 (Power)

Pin 1 - VCC "+"
Pin 2 - GND "-"

Header p3 (UART/Serial)

Pin 1 - GND "-"
Pin 2 - VCC "+"
Pin 3 - RXD "R"
Pin 4 - TXD "T"

Example Design

This example design uses 2 AAA batteries to power a TGAM module. Both the EEG and REF electrodes are connected with shielded wires to TGAM, and both of the shielded ground wires are connected to the UART ports of a processor.

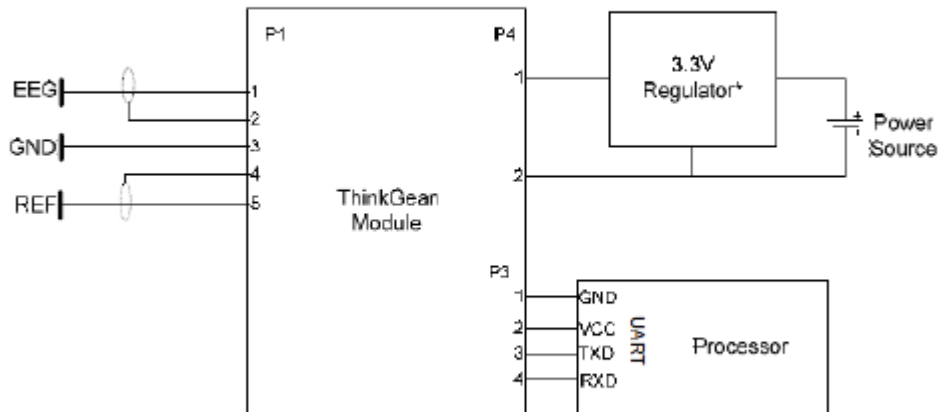


Fig 4 Example Schematic * Normal Output mode Includes the following output: poor quality value, EEG value, attention value, and meditation value

Sensor AD8232

FEATURES

- Fully integrated single-lead ECG front end
- Low supply current: 170 μ A (typical)
- Common-mode rejection ratio: 80 dB (dc to 60 Hz)
- Two or three electrode configurations
- High signal gain ($G = 100$) with dc blocking capabilities
- 2-pole adjustable high-pass filter
- Accepts up to ± 300 mV of half cell potential
- Fast restore feature improves filter settling
- Uncommitted op amp
- 3-pole adjustable low-pass filter with adjustable gain
- Leads off detection: ac or dc options
- Integrated right leg drive (RLD) amplifier
- Single-supply operation: 2.0 V to 3.5 V
- Integrated reference buffer generates virtual ground
- Rail-to-rail output
- Internal RFI filter
- 8 kV HBM ESD rating
- Shutdown pin
- 20-lead 4 mm \times 4 mm LFCSP package

APPLICATIONS

- Fitness and activity heart rate monitors
- Portable ECG
- Remote health monitors
- Gaming peripherals
- Biopotential signal acquisition

FUNCTIONAL BLOCK DIAGRAM

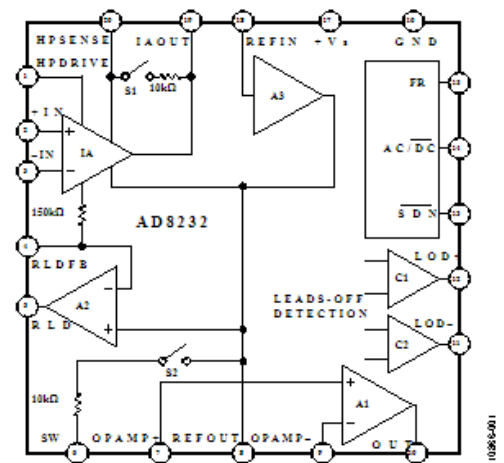


Figure 1.