

Implementación del algoritmo K-means para clusterización de señales EEG durante la aplicación de una prueba Stroop

Implementation of the K-means algorithm for clustering of EEG signals during the application of a Stroop test

Paúl Esteban Cárdenas¹ <https://orcid.org/0000-0003-0526-8519>,
Daniela Prado¹ <https://orcid.org/0000-0003-1241-1782>, Bruno Iglesias², Ronnie Urdiales², Marcos Orellana² <https://orcid.org/0000-0002-3671-9362>, Priscila Cedillo^{1, 2} <https://orcid.org/0000-0002-6787-0655>

¹Universidad de Cuenca, Cuenca, Ecuador
paul.cardenas@ucuenca.edu.ec, daniela.pradoc@ucuenca.edu.ec,
priscila.cedillo@ucuenca.edu.ec

²Universidad del Azuay, Cuenca, Ecuador
biglesias@es.uazuay.edu.ec, rurdiales@es.uazuay.edu.ec,
marore@uazuay.edu.ec, icedillo@uazuay.edu.ec



Esta obra está bajo una licencia internacional
Creative Commons Atribución-NoComercial 4.0.

Enviado: 2021/07/11
Aceptado: 2021/09/28
Publicado: 2021/11/30

Resumen

El análisis de datos y la generación de modelos mediante aprendizaje automático (Machine Learning - ML por sus siglas en inglés) es una de las técnicas más usadas por la comunidad científica para obtener conocimiento que no se puede interpretar o analizar a simple vista. En este documento, se aborda el tema específico de la obtención, procesamiento de datos y modelado de los mismos mediante técnicas de agrupamiento, concretamente el algoritmo de K-Means. Los datos se obtienen a través de una herramienta de interfaz cerebro computador (ICC - BCI por sus siglas en inglés) basada en electroencefalograma (EEG), mientras un sujeto de prueba realiza una tarea de tipo Stroop, la cual permite a los psicólogos evaluar el control inhibitorio de las personas. La aplicación de este tipo de algoritmo en esta área específica, busca poder identificar patrones dentro de las señales EEG relacionados al estado del sujeto. La captura de datos se ejecuta en tres diferentes horas del día, matutina, vespertina, y nocturna, durante dos días consecutivos, con el objetivo de obtener variabilidad en los datos. Si bien la muestra de datos es pequeña, se puede emplear como punto de partida para el análisis del algoritmo K-means en señales EEG durante la ejecución del test Stroop. Los resultados se muestran tanto desde el punto de vista técnico como psicológico, y se puede ver en la

Sumario: Introducción, Trabajos relacionados, Metodología, Resultados, Discusión y Conclusiones.

Como citar: Cárdenas, P., Prado, D., Iglesias, B., Urdiales, R., Orellana, M. & Cedillo, P. (2021). Implementación del algoritmo K-means para clusterización de señales EEG durante la aplicación de una prueba Stroop. *Revista Tecnológica - Espol*, 33(2), 172-188.
<http://www.rte.espol.edu.ec/index.php/tecnologica/article/view/847>

clusterización realizada con las señales en el dominio del tiempo, que existe un patrón de agrupación según la hora del día en la que se realiza el test. Por otra parte, para la clusterización con las señales en el dominio de la frecuencia este patrón no resulta tan evidente. Desde la perspectiva psicológica, se comprueba que el proceso de aprendizaje y de acomodación al momento de realizar una prueba psicológica, disminuye su potencial.

Palabras clave: aprendizaje automático, electroencefalograma, agrupamiento, evaluación neuropsicológica, Interfaz Cerebro-Computador.

Abstract

Data analysis and model generation through machine learning (Machine Learning - ML) is one of the techniques most used by the scientific community to obtain knowledge that cannot be interpreted or analyzed with the naked eye. In this document, the specific issue of data collection, processing, and modeling is addressed using clustering techniques, specifically the K-Means algorithm. The data were obtained through a brain-computer interface (ICC - BCI) tool based on electroencephalogram (EEG), while a test subject performed a Stroop-type task, which allows psychologists to evaluate people's inhibitory control. The application of this type of algorithm in this specific area seeks to identify patterns within the EEG signals related to the subject's state. Data capture was carried out at three different hours of the day, morning, evening, and night, for two consecutive days, to obtain variability in the data. Although the data sample is small, it will serve as a starting point for analyzing the K-means algorithm in EEG signals during the execution of the Stroop test. The results obtained are shown both from a technical and psychological point of view. It can be seen in the clustering carried out with the signals in the time domain that there is a clustering pattern according to the time of day in which the test was performed. On the other hand, this pattern is not so evident for the clustering with the signals in the frequency domain. From the psychological perspective, it is found that the learning and accommodation process at the time of performing a psychological test decreases its potential.

Keywords: machine learning, electroencephalographic, groping, neuropsychological assessment, Brain-Computer Interface.

Introducción

Desde hace algunos años, la Inteligencia Artificial (IA) es una de las de las herramientas más importantes para el desarrollo de nuevas tecnologías que ayudan a las personas en tareas que, por dificultades físicas o psicológicas endógenas, tienen limitaciones al momento de ejecutarlas (Russell & Norvig, 2004). Debido a la proliferación de la IA, esta se ha dividido en varios subcampos para cubrir la diversidad de aplicaciones que presenta (Ayodele, 2010). Entre ellos está el desarrollo de técnicas de *Machine Learning* (ML), que es uno de los campos más importantes hoy en día, ya que, se basa en la aplicación de técnicas estadísticas como regresiones lineales, regresiones logísticas, entre otras, para la creación de algoritmos mucho más eficientes, con el objetivo de identificar patrones o realizar predicciones sobre un conjunto de datos.

En este contexto, se encuentran dos clasificaciones: algoritmos supervisados y no supervisados. En ambos casos, los algoritmos reciben variables de entrada o *inputs* y pueden o no contener variables de salida. Es decir, la diferencia radica en la presencia o no de variables de salida o *target*. Precisamente en el caso de los algoritmos supervisados se requiere indicar este tipo de variable; mientras que en los algoritmos no supervisados, no se requiere de esta variable. Por lo que este tipo de algoritmos buscan encontrar patrones que permitan entender

el comportamiento de los datos de mejor manera, ya sea con el fin de agruparlos, clasificarlos o realizar predicciones basadas en datos (Nasteski, 2017).

El *Machine Learning* se utiliza en un amplio espectro de contextos (Carvalho et al., 2019), uno de los más relevantes es la aplicación en el área de la salud, y particularmente en el cuidado del paciente (Ben-Israel et al., 2020). Además, en este subcampo de la IA se emplea para el análisis de información del comportamiento humano que sirve como insumo para el desarrollo de modelos de entrenamiento cognitivo (Rennie et al., 2020). Es decir, el *Machine Learning* brinda soporte para el análisis de datos provenientes de la intervención neuropsicológica para el entrenamiento personalizado de las funciones cognitivas (Shani et al., 2019), entendidas como los procesos mentales involucrados en la adquisición de conocimiento, la manipulación de la información y el razonamiento (Montañés & de Brigard, 2011).

Existen distintas funciones cognitivas como la atención, concentración, memoria, lenguaje, razonamiento, control inhibitorio, entre otras (Trápaga et al., 2018). Particularmente, el control inhibitorio es la capacidad que tiene un individuo para controlar reacciones impulsivas tanto a nivel atencional como comportamental, esta función depende principalmente de la corteza pre frontal (Peña Barrientos et al., 2017). A pesar de que Scarpina y Tagini (2017) realizan varias pruebas neuropsicológicas para analizar esta función a lo largo de los años, la técnica predilecta por expertos, y que se mantiene desde el siglo pasado, es el test de Stroop (Stroop, 1935).

Este test atencional tiene la finalidad de tasar la capacidad de un sujeto para resistir la interferencia de tipo verbal. Se trata de una herramienta de tamizaje útil para el diagnóstico de patologías del desarrollo y adquiridas (por ejemplo: Trastorno del espectro autista, traumatismos craneoencefálicos, síndrome de Tourette, trastorno por déficit de atención con hiperactividad). También se emplea para identificar la capacidad del sujeto de dar una respuesta impidiendo la intrusión de un proceso automático, donde la atención selectiva juega un papel fundamental, ya que el ser humano requiere controlar su capacidad atencional orientado a los estímulos de mayor beneficio y adaptación. Esta tarea consiste en indicar el color de la tinta con la que está escrita una palabra que no coincide con su significado, por ejemplo, si se escribe la palabra “azul” aunque está impresa con color verde, produce una interferencia semántica en el cerebro que causa un aumento en la latencia de respuesta verbal del sujeto de prueba.

A partir de la explicación dada, el objetivo de esta investigación se centra en implementar un modelo que agrupe las señales del Electroencefalograma (EEG) que se obtienen durante la ejecución de la prueba Stroop de un sujeto de ensayo con el fin de identificar patrones o características del estado actual del individuo durante la aplicación del test Stroop. Para ello, se plantea el empleo de un algoritmo muy utilizado dentro del *Machine Learning* denominado *K-Means* (Likas et al., 2003), el cual es un algoritmo de aprendizaje no supervisado que pertenece al tipo de algoritmo de agrupamiento o también conocido como *Clustering* (Alqurashi & Wang, 2019). Está claro que dentro del *Machine Learning* existen modelos o algoritmos más potentes que se pueden utilizar en este contexto, sin embargo, se considera este algoritmo debido a que es relativamente de fácil implementación y como una primera aproximación al objetivo de identificar patrones pertenecientes al control inhibitorio en señales EEG.

Este trabajo se organiza de la siguiente manera: la Sección 2 presenta estudios relacionados, la Sección 3 introduce la solución y método planteado, la Sección IV explica el caso de estudio donde se aplica la solución y el resultado del análisis de datos, finalmente se expone la conclusión.

Trabajos relacionados

Las formas de evaluación, diagnóstico y predicción del estado de las funciones cognitivas, y particularmente, del control inhibitorio, avanzan de manera acelerada los últimos años (Emerenziani et al., 2020). Gracias a esto, se combinan metodologías de investigación de las ciencias sociales, médicas y técnicas; con el objetivo de que los estudios sean más exactos y que varias disciplinas se puedan juntar en una misma investigación para comprender, de mejor manera, los fenómenos humanos (Onwuegbuzie & Leech, 2005).

Priya et al. (2020) realizan un estudio en el que se predicen variables de condiciones psicológicas como la depresión, ansiedad y estrés mediante *Machine Learning*, específicamente árboles de decisión y algoritmos de Naive-Bayes. Ellos no consideran dividir los factores psicológicos en niveles o tipos, solo detectan la presencia o ausencia de alteraciones cognitivas. Se puede resaltar que los autores consideran que, en algunas ocasiones, los algoritmos de *Machine Learning* dan falsos positivos o falsos negativos, frente a ello, desarrollan un algoritmo que detecta estos falsos resultados, haciendo que el sistema sea mucho más eficiente.

Por otra parte, Wang et al. (2018) desarrollan un sistema para detectar la ansiedad usando un casco para la interacción cerebro computador (*Brain Computer Interface – BCI*) mediante la captura de señales EEG del sujeto. Medina et al. (2018) por su parte, depuran las señales con técnicas específicas y; posteriormente, desarrollan un sistema que detecta el nivel de atención del sujeto evaluado y predice cuantitativamente, por los diferentes picos que existen en determinadas señales, si el sujeto evaluado tiene ansiedad; esto resulta útil porque hay muy poco margen de error utilizando estas técnicas.

Kaneta et al. (2016) investigan la influencia del efecto Stroop en la onda cerebral, específicamente en los potenciales relacionados con eventos (ERP – por sus siglas en ingles). Para ello hace uso de la técnica de agrupamiento máquina de vector de soporte (*Super Vector Machine – SVM*), mediante la cual clasifican a las señales en dos grupos dependiendo si el estímulo es dirigido o no dirigido.

How et al. (2015), realizan la implementación de un clasificador mediante SVM, con el objetivo de clasificar los niveles de estrés en las señales EEG. Para esto utilizan la prueba Stroop como factor estresante. En sus resultados logran clasificar en cuatro niveles de estrés con una precisión media del 67,06%, en 3 niveles con una precisión del 75,22% y en dos niveles de estrés con una precisión del 85,71%.

Lim et al. (2016) analizan el estrés cognitivo en las señales EEG de 25 sujetos durante la aplicación del test Stroop. Para esto aplican tres técnicas o algoritmos como son: redes neuronales artificiales, k-Nearest Neighbor (KNN – por sus siglas en ingles) and Linear Discriminant Analysis (LDA – por sus siglas en ingles). Sus resultados les permiten identificar que el algoritmo que tiene mejor rendimiento es el KNN.

Como se puede observar, se realizan diferentes estudios respecto al análisis de señales EEG durante la aplicación del test Stroop, con el objetivo de clasificar o predecir distintas condiciones psicológicas (Al-Nafjan et al., 2017; Priya et al., 2020). Sin embargo, muchos de ellos se centran en la detección del estrés, siendo necesario investigar otros estados mentales o funciones cognitivas que puedan producirse durante la aplicación de este test (Popov et al., 2018). A partir de una indagación previa, se comprueba que no existe una investigación centrada en el análisis de señales EEG para determinar patrones del estado mental que se puedan relacionar al control inhibitorio. Por lo tanto, este artículo se centra en la

implementación del algoritmo *K-Means*, con el objetivo de que, mediante la toma de datos de las señales EEG de un sujeto durante la ejecución del test Stroop, se genere un agrupamiento que permita identificar patrones en las señales EEG capturadas en las diferentes fases del test y sesiones realizadas, para contrastarlos con los datos tomados de forma manual.

Metodología

La metodología de investigación seleccionada para la presentación de esta investigación es la propuesta por Runeson et al. (2012) que implica la delimitación de estudios de caso. A continuación se explican los pasos empleados: A) diseño y planificación, B) consideraciones éticas, C) preparación y recolección de datos, D) análisis e interpretación de los resultados, E) amenazas a la validez. Sin embargo, en este apartado a fines metodológicos se presentan los pasos A, B y C; dejando los pasos D y E para la sección 4. Resultados.

Diseño y planificación

El objetivo central de este estudio se presenta por medio de la propuesta *Goal Question Metric* (GQM) de Basili et al. (1994), ya que se puede plasmar de manera objetiva lo que se requiere evaluar por medio del estudio de caso (Tabla 1).

Tabla 1

Aproximación GQM

Analizar:	Patrones de las señales EEG relacionados al estado actual del sujeto durante la ejecución de la tarea tipo Stroop verbal.
Con el propósito:	Generar clusters de los patrones encontrados que proporcionen información de utilidad clínica.
Con respecto a:	Sujeto de estudio adulto sin precedentes de neuropatología.
Desde el punto de vista de:	Técnico.

Consideraciones éticas

Aunque un estudio de investigación se basa principalmente en la confianza entre el investigador y el caso (Amschler & Pradhan, 2001), se consideran algunas medidas de prevención explícitas para evitar problemas futuros relacionados con el bienestar del participante. Los principales factores éticos incluidos en este estudio de caso son: i) consentimiento informado, ii) revisión de la aprobación de un comité de bioética, iii) confidencialidad y iv) retroalimentación de los hallazgos.

Preparación y recolección de datos

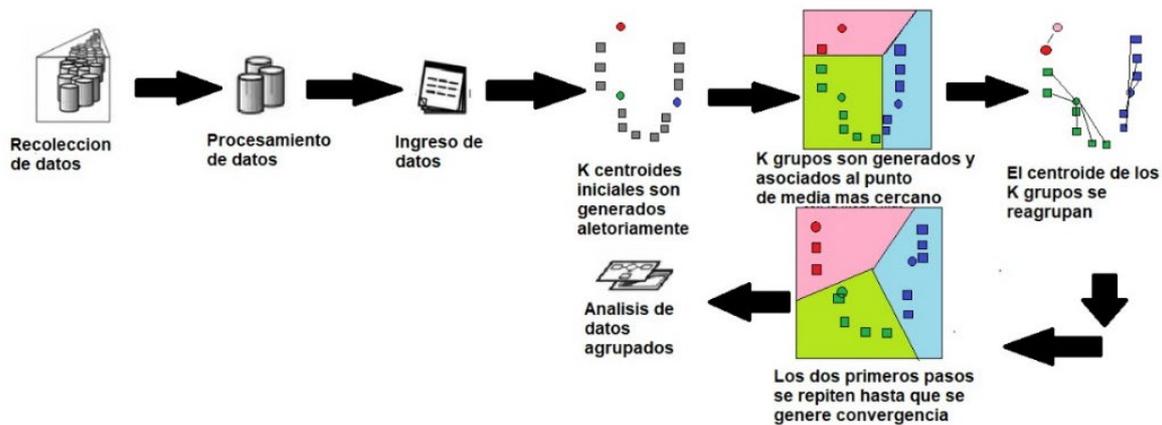
A continuación, se explica la metodología que orienta a la obtención de datos para esta investigación, esta corresponde a la propuesta de Fayyad et al. (1996) titulada *Knowledge Discovery in Databases* (KDD), que se refiere al proceso general de descubrimiento de conocimiento útil a partir de datos. Para garantizar que se deriven conclusiones útiles, se efectúan pasos como la preparación, selección y limpieza de datos. Además de la incorporación de conocimientos apropiados y la interpretación adecuada de los resultados de la minería.

En la Figura 1 se presentan los pasos para la recolección de los datos, clusterización a partir de la aplicación del algoritmo *K-Means* y análisis de los clusters encontrados. La ventaja de utilizar este procesamiento, radica en que este es un proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles en los datos.

Los datos obtenidos del EEG, con los que se entrenó el modelo, se recolectaron mediante un casco BCI basado en EEG. La captura de los datos se la realizó en un solo sujeto mientras ejecutaba la prueba Stroop, en seis sesiones diferentes distribuidas en dos días, en tres periodos de tiempo: al despertar, al medio día y antes de dormir.

Figura 1

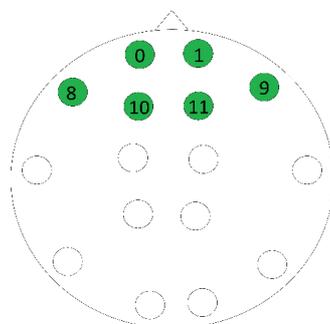
Pasos que constituyen el proceso KDD (Fayyad et al., 1996)



Específicamente, se utilizó un casco OpenBCI, con una placa “Cyton Boisenring Board + Daisy”, la cual permitió trabajar con 16 canales con electrodos secos no invasivos. Para visualizar los datos en tiempo real se utilizó la interfaz de usuario del software de OpenBCI (Suryotrisongko & Samopa, 2015), la cual permitió guardar el archivo de datos tanto en formato de texto (.txt), como en formato europeo de 24 bits (.bdf). Para este estudio, se utilizó el archivo de texto, ya que permitió analizar de forma simple las señales para los objetivos planteados. El archivo se organizó en columnas que corresponden a cada uno de los canales utilizados para el registro, en este caso se emplearon 6 canales posicionados en la corteza prefrontal, correspondientes a los electrodos 0, 1, 8, 9, 10, y 11, como se puede observar en la Figura 2. Esta selección de canales se la realizó, ya que es en la corteza prefrontal donde se podría evidenciar con mayor claridad la actividad de control inhibitorio en las señales cerebrales del sujeto. Por otra parte, en las filas se mostraron los potenciales en micro voltios de cada uno de los registros realizados en el tiempo, con una frecuencia de 125 Hz.

Figura 2

Distribución de los electrodos en la corteza cerebral

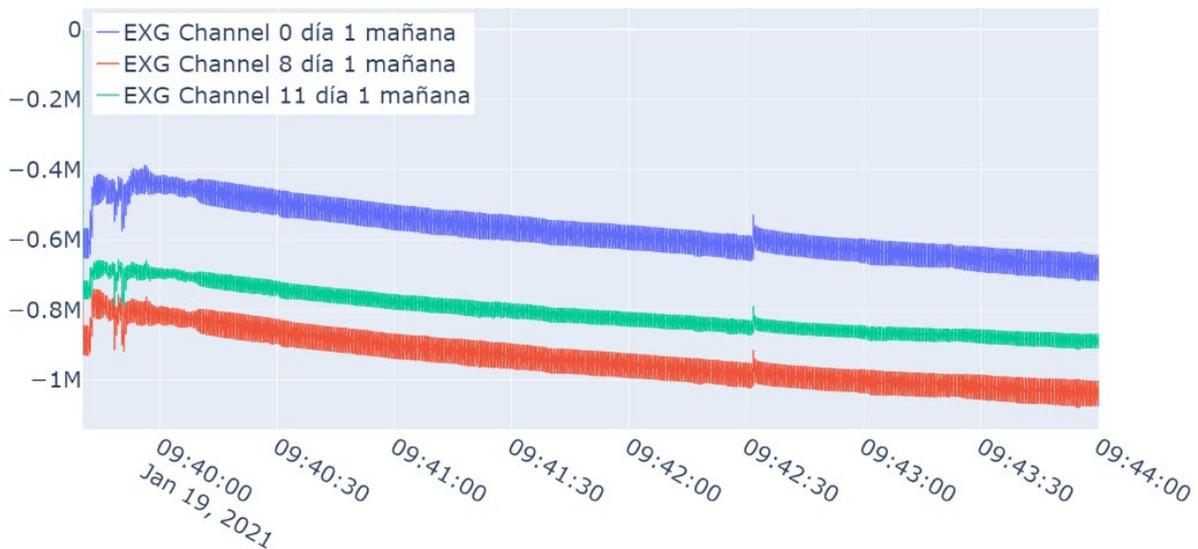


Para el preprocesamiento de los datos se emplearon diferentes librerías en Python, como Pandas, Numpy, Plotly, entre otras. Para utilizar el conjunto de datos en un modelo como *K-Means*, primero fue necesario realizar un análisis y preprocesamiento, lo cual permitió

desarrollar una correcta extracción de características de las señales capturadas. Para este estudio se analizaron las señales EEG tanto en el dominio del tiempo como de la frecuencia, lo que dio paso a la identificación de la existencia de frecuencias no deseadas. El análisis se lo efectuó para cada uno de los canales de las seis sesiones realizadas. En la Figura 3 se puede observar un ejemplo de las señales captada por los seis canales durante la sesión de la tarde del día uno en el dominio del tiempo.

Figura 3

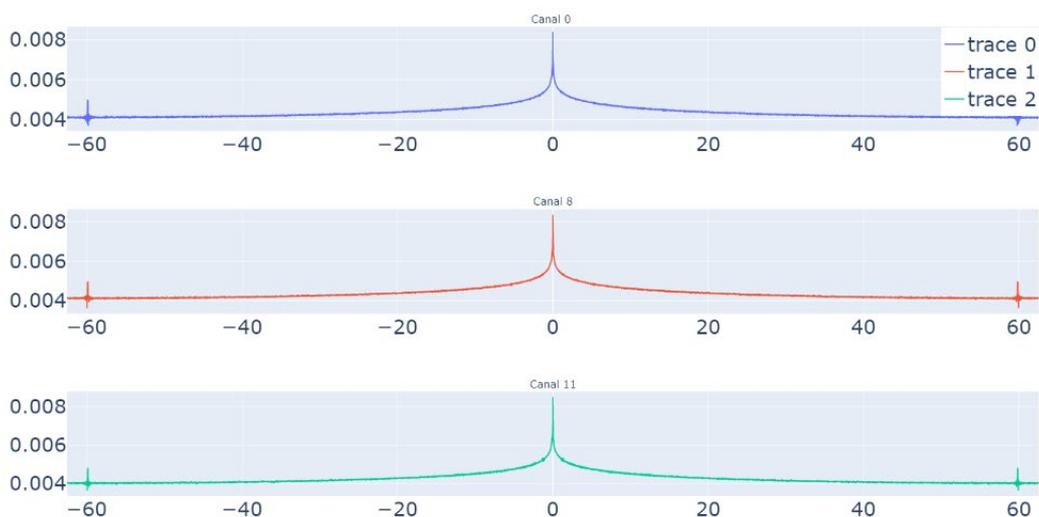
Señales en el dominio del tiempo de la sesión de la tarde en el día 1



Para realizar el análisis en el dominio de la frecuencia, se utilizó la transformada rápida de *Fourier* (FFT). La Figura 4 muestra un ejemplo de las señales captada por los seis canales durante la sesión de la tarde del día uno en el dominio de la frecuencia.

Figura 4

Señales EEG en el dominio de la frecuencia

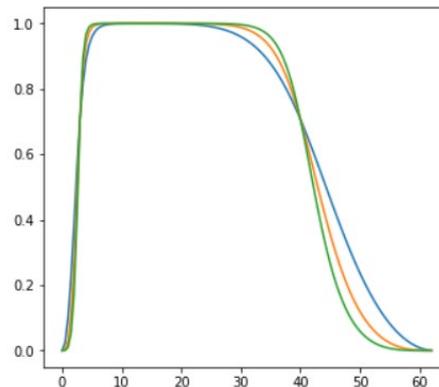


Después del análisis realizado, como se puede ver en la Figura 4, el espectro de frecuencia de las señales presentó un pico en frecuencias bajas cercanas a cero Hz, por lo que se aplicó un filtro paso banda entre tres y 40 Hz. Como se puede ver en la Figura 5, con el fin

de eliminar frecuencias atípicas que puedan representar ruido o niveles muy bajos de frecuencia correspondientes a ondas delta, y a su vez tener una mejor visualización de las bandas Theta, Alpha y Beta, ya que estas son las que contendrían mayor información respecto al control inhibitorio en la actividad cerebral del sujeto.

Figura 5

Filtro pasa banda aplicado a las señales EEG



En la Figura 6, se puede ver un ejemplo de tres de las señales capturadas por los seis canales durante la sesión de la tarde del día uno, filtradas en el dominio de la frecuencia y del tiempo. Además, se han atenuado las frecuencias cercanas a cero y se puede visualizar de mejor manera las frecuencias superiores correspondientes a las bandas de interés antes mencionadas.

Figura 6

Señales filtradas en el dominio de la frecuencia y tiempo



Para la implementación del algoritmo *K-Means*, se realizaron dos pruebas en el primer caso, con los datos en el dominio del tiempo, y en el segundo caso en el dominio de la frecuencia. En el primer caso, las características o columnas a utilizar, fueron de cada uno de los canales seleccionados previamente, y las filas cada una de las muestras tomadas a lo largo del tiempo, concatenando una sesión tras otra en orden cronológico. Por otra parte, en el segundo caso, para reducir la dimensionalidad de los datos en el dominio de la frecuencia, se realizó el cálculo y selección de características principales de las señales EEG capturadas. Entre las principales características utilizadas en este caso, se hizo uso de la densidad espectral de potencia, densidad relativa de potencia, potencia máxima de frecuencia, potencia mínima de

frecuencia, media de la potencia de la frecuencia, y finalmente la varianza y covarianza de la frecuencia.

Una vez que se obtuvo la tabla final con los datos procesados, tanto en el dominio del tiempo como de la frecuencia, se realizó la normalización de los valores numéricos, para así obtener un mejor desempeño del algoritmo. En este estudio, debido a la poca cantidad de datos recabados, se utilizó el 100% del set de datos para realizar el agrupamiento mediante el algoritmo de *K-Means*.

Como ya se mencionó anteriormente, este estudio se enfoca en las bandas Theta, Alpha, y Beta, en la Figura 7 se puede observar un ejemplo de cómo están distribuidas las bandas de frecuencias en las señales ya filtradas que se utilizaron para entrenar el algoritmo. De igual manera en la Figura 8 se puede ver la densidad espectral de potencia de uno de los canales de las señales EEG a ser utilizadas para entrenar el algoritmo, en azul se puede ver las frecuencias a ser utilizadas para este estudio.

Figura 7

Bandas de frecuencias de las señales a ser utilizadas para entrenar el algoritmo

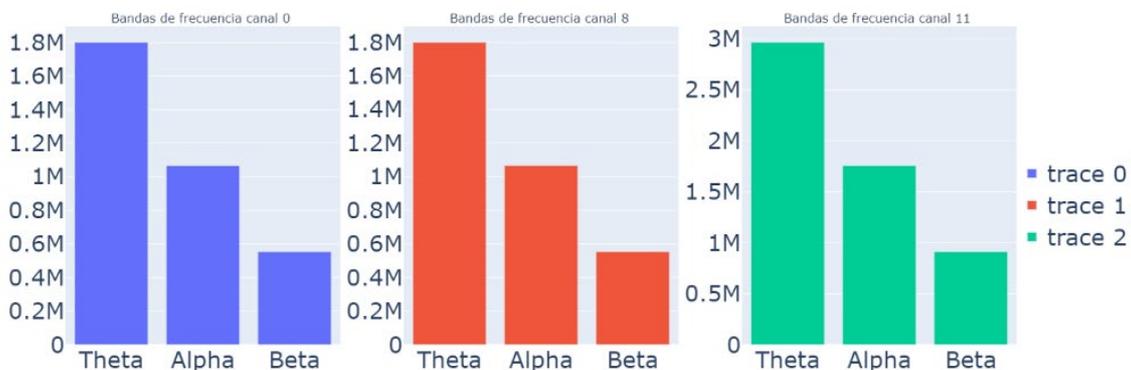
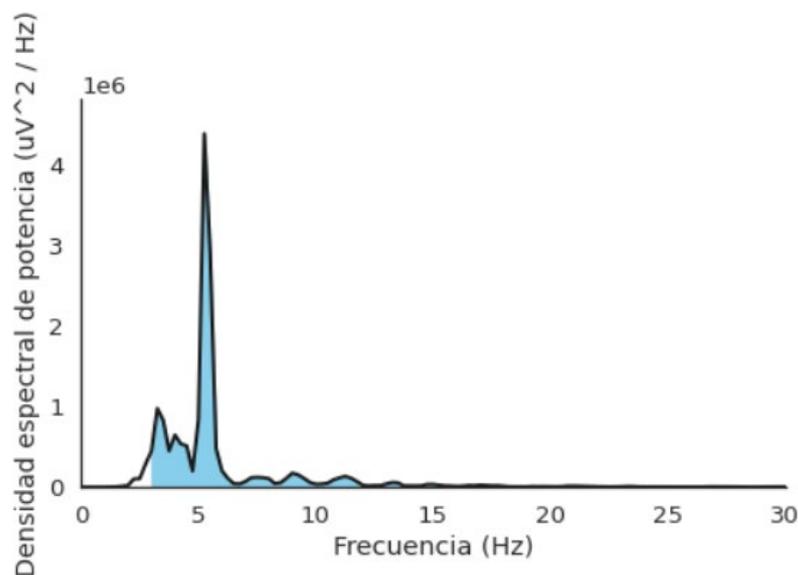


Figura 8

Densidad espectral de potencia de las señales a ser utilizadas para entrenar el algoritmo



Resultados

A continuación, se presenta la implementación del algoritmo *K-means* y el análisis de datos agrupados y las amenazas a la validez.

Análisis de datos agrupados

Para ejecutar el algoritmo de *Clustering*, se utilizó la librería *Sklearn* de *Python*, por razones de afinidad tanto con la librería como con el lenguaje de programación.

El algoritmo, en un primer caso, recibió como entrada el conjunto de datos de EEG, en este proceso, los seis canales en el dominio del tiempo seleccionados y preprocesados previamente. Como métrica de evaluación y para una elección adecuada del número de clústeres a realizar, se utilizó el método del codo, el cual empleó la distancia media de las observaciones a su centroide, es decir, se fijó en las distancias intra clúster, si el número de clústeres se incrementa, la varianza intra clúster tiende a disminuir. De acuerdo a la gráfica obtenida, como se puede ver en la Figura 9, se tiene que el número ideal de clústeres está entre dos a seis, para el caso de las señales en el dominio del tiempo.

Figura 9

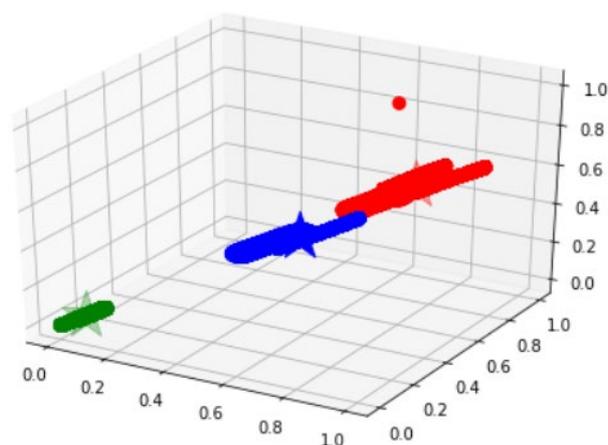
Gráfica de codo para los datos en el dominio del tiempo



Después de varias pruebas y análisis realizadas con diferentes valores de k clústeres, se estableció que tres es el número adecuado de clústeres. En la Figura 10 se puede ver el resultado de la clusterización realizada en tres dimensiones.

Figura 10

Clusterización con las señales en el dominio del tiempo



En un segundo caso, el algoritmo recibió como entrada los datos de las características principales extraídas de los canales en el dominio de la frecuencia seleccionados y preprocesados, mencionados previamente. Al igual que en el primer caso, como métrica de evaluación y para una elección adecuada del número de clústeres a realizar, se utilizó el método del codo. De acuerdo a la gráfica obtenida, como se puede ver en la Figura 11, se tiene que el número ideal de clústeres está entre dos y ocho, para el caso de las señales en el dominio de la frecuencia. Esto sucede posiblemente, porque en este dominio, se pudo detectar algunas señales en el espectro de frecuencia atípicas con respecto al resto de canales, principalmente de la sesión del día uno en la mañana. Otro aspecto importante a mencionar, es que el número de muestras disminuyó a 36, ya que al realizar el proceso de extracción de características en el dominio de la frecuencia, cada una de las señales capturadas correspondían a una muestra, mientras que las características extraídas a las dimensiones. Esto por un lado ayudó a que el rendimiento del algoritmo sea mejor, sin embargo, la poca cantidad de muestras utilizadas pueden representar un inconveniente para la obtención de resultados adecuados.

Figura 11

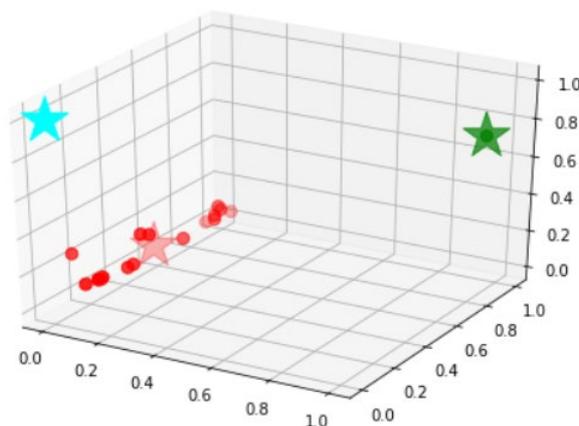
Gráfica de codo para los datos en el dominio de la frecuencia



Después de varias pruebas y análisis realizados con diferentes valores de k clústeres, se estableció que el número ideal de clústeres, con el fin de identificar similitudes con la clusterización realizada en el dominio del tiempo. En la Figura 12 se puede ver el resultado de la clusterización elaborada en tres dimensiones.

Figura 12

Clusterización con las señales en el dominio de la frecuencia



Los resultados obtenidos después de la clusterización se analizaron con base en tres características: i) hora del día (mañana, tarde, noche), ii) día (día 1, día 2), e ii) interferencia (baja, media, alta), esta última establecida según los valores calculados más adelante, obtenidos en la realización del test.

De acuerdo a esto en la Tabla 2 se puede observar la distribución de los resultados obtenidos mediante los dos casos de clusterización.

De acuerdo con estos resultados, se puede evidenciar que en la clusterización, en el dominio del tiempo, existe una mejor distribución de los datos, ya que se logra distinguir claramente que el algoritmo ha separado a las señales casi de manera perfecta según la hora del día en la que fueron capturadas, sin embargo, en relación tanto al día como a la interferencia, no se encontraron mayores diferencias entre clústeres ya que los porcentajes se distribuyen de manera similar.

En cuanto a la segunda clusterización, en el dominio de la frecuencia, los resultados no evidencian una clara distribución de los datos, ya que los porcentajes son similares para cada clúster, esto puede deberse a la poca cantidad de datos obtenidos, se puede probar a futuro el algoritmo en este dominio con un espectro mucho más amplio de datos y con más sujetos de estudio.

Tabla 2

Porcentaje de la distribución de los datos en los clústeres realizados a base de tres características analizadas

CLUSTERIZACIÓN EN EL DOMINIO DEL TIEMPO				CLUSTERIZACIÓN EN EL DOMINIO DE LA FRECUENCIA			
JORNADA	CLUSTER 1 (ROJO)	CLUSTER 2 (VERDE)	CLUSTER 3 (CELESTE)	JORNADA	CLUSTER 1 (ROJO)	CLUSTER 2 (VERDE)	CLUSTER 3 (CELESTE)
Matutina	34,56%	0,00%	0,00%	Matutina	27,78%	2,78%	2,78%
Vespertina	17,72%	14,73%	0,00%	Vespertina	27,78%	0,00%	5,56%
Nocturna	0,00%	0,00%	32,98%	Nocturna	27,78%	0,00%	5,56%
DÍA	CLUSTER 1 (ROJO)	CLUSTER 2 (VERDE)	CLUSTER 3 (CELESTE)	DÍA	CLUSTER 1 (ROJO)	CLUSTER 2 (VERDE)	CLUSTER 3 (CELESTE)
1	37,03%	0,00%	18,45%	1	41,67%	2,78%	5,56%
2	15,25%	14,73%	14,54%	2	41,67%	0,00%	8,33%
INTERF.	CLUSTER 1 (ROJO)	CLUSTER 2 (VERDE)	CLUSTER 3 (CELESTE)	INTERF.	CLUSTER 1 (ROJO)	CLUSTER 2 (VERDE)	CLUSTER 3 (CELESTE)
Baja	0,00%	0,00%	18,45%	Baja	41,67%	0,00%	8,33%
Media	327,03%	0,00%	0,00%	Media	27,78%	2,78%	2,78%
Alta	15,25%	14,73%	14,54%	Alta	13,89%	0,00%	2,78%

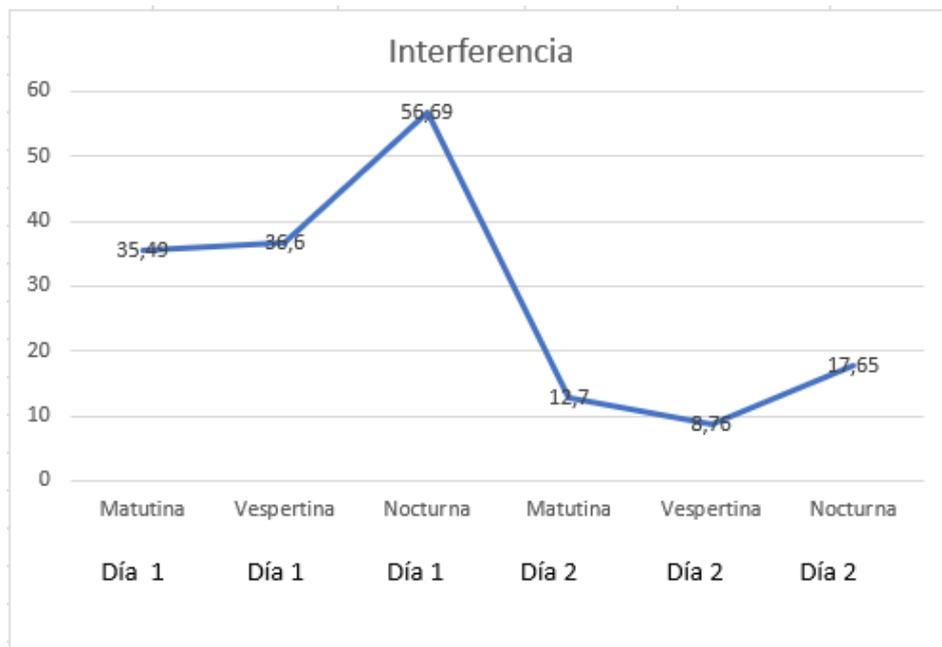
La Figura 13, presenta los resultados calculados según indica el manual de Scarpina & Tagini (2017), para ello, se emplearon los cálculos presentados en la Figura 14 (Stroop, 1935).

Como se ve en la Figura 14, existen cinco variables que se utilizan para obtener el valor de la interferencia, las cuales son:

- P: número de palabras leídas en la primera fase del test
- C: número de colores leídos en la segunda fase del test
- PC: número de palabras leídas en la tercera fase del test
- PC': Valor obtenido de la división entre la suma y el producto del valor de P con el valor de C
- Interferencia: Valor obtenido de la diferencia entre el valor de PC y el valor de PC'

Figura 13

Comparación de la interferencia en los días de prueba

**Figura 14**

Fórmula utilizada para el cálculo de la interferencia al realizar la prueba de tipo Stroop (Stroop, 1992)

	PD	PT
P		
C		
PC		
$\frac{P \times C}{P + C} = PC'$		
$PC - PC' = \text{INTERF.}$		

Amenazas a la validez

En los procesos de evaluación empírica existen cuatro tipos de amenazas a la validez que son necesarios señalar: i) validez de constructo, ii) validez interna, iii) validez externa y iv) fiabilidad.

Validez de constructo

La validez de constructo se centra en la relación entre la teoría detrás del caso de estudio y la observación. Para ello, se considera dos propuestas teóricas validadas. Por un lado el efecto

Stroop (1935) que ha sido validado estadísticamente por la comunidad científica para la valoración del estado de control inhibitorio y atención selectiva; además, el algoritmo *k-means* que ha sido demostrado matemáticamente (Xin, 2011).

Validez interna

La validez interna está relacionada con la calidad de los resultados del estudio. En este estudio de caso, se cuenta con información limitada para el análisis; por lo tanto, el resultado de la clusterización puede presentar imprecisiones; por ello, se procede a calificar los resultados de manera manual, para contrastar la información.

Validez externa

La validez externa está relacionada a la capacidad de generalización de los resultados obtenidos. En este estudio de caso, al contar con información de un solo sujeto, los resultados son poco generalizables. Sin embargo, el estudio sirve como una primera aproximación a la aplicación de tecnología a este tipo de herramientas psicológicas, puntualmente algoritmos de machine learning que permitan identificar patrones en los datos, y sirvan como soporte a la intervención de un psicólogo o personal de la salud.

Fiabilidad

El test Stroop tiene un alfa de Cronbach de 0.76, lo que significa que es un instrumento confiable para medir el control inhibitorio (Eliorraga, 2017).

Discusión

En este apartado, se presenta la discusión desde dos puntos de vista; por un lado, el técnico que visibiliza las fortalezas y limitaciones del estudio; por otro, desde el punto de vista psicológico y las implicaciones clínicas que este tipo de herramientas tienen para los profesionales de la salud cognitiva.

Desde el punto de vista técnico

La aplicación de algoritmos de agrupamiento a señales EEG ayudan a identificar patrones y relaciones con ciertas tareas que realiza un sujeto, como en este caso el teste Stroop. En este sentido el algoritmo implementado permite ver ciertas relaciones en los clústeres obtenidos con las señales en el dominio del tiempo, respecto a los horarios de realización del test, de esta forma se puede plantear la interrogante si el horario en el que un sujeto realiza el test influye en sus resultados o en el comportamiento de sus ondas de la corteza prefrontal. Por otro lado, la clusterización con las señales en el dominio de la frecuencia no muestra resultados del todo óptimos, esto debido a la poca cantidad de datos recabados para analizar las señales en este dominio, que por lo general permite obtener mayor información de las señales que en el dominio del tiempo.

Sin embargo, el obtener los modelos en este estudio, es un aporte a las pruebas del procesamiento de las señales y la extracción de características de las mismas, así como identificar los factores que influyen en la existencia de un modelo de clusterización en este dominio. Esto a su vez, da la pauta para una siguiente investigación, en la que se prueben los modelos planteados con un conjunto de datos mucho más robusto.

Desde el punto de vista psicológico

Los resultados que se derivan de la aplicación del test Stroop a un sujeto sin patología severa, denota que, en el segundo día, tal y como lo muestra la Figura 13, presenta un proceso de adecuación y aprendizaje de la tarea, por lo que, el efecto Stroop pierde fuerza. Además, se puede observar cómo la interferencia es mayor en el horario nocturno, es decir, se comprueba

que la capacidad del ser humano para inhibir o controlar las respuestas impulsivas (o automáticas), y generar respuestas mediadas por la atención y el razonamiento se reduce cuando existen factores como el cansancio o fatiga de la jornada laboral (MacLeod, 1991; Stroop, 1935). Todo esto, se comprueba con los hallazgos representados en la Tabla 1, particularmente en la clusterización. En el dominio del tiempo brinda información relevante, se observa que, el clúster uno concentra datos que corresponden a la jornada matutina, mientras que en el tres está la vespertina.

Tomando en cuenta que la media de interferencia para personas de entre 16 a 44 años oscila desde 30 a -30 (Stroop, 1935), en el día uno, el sujeto supera la media en interferencia, y que en el día dos se encuentra justo a la mitad. Cabe recalcar que, mientras menos interferencia presente el sujeto al realizar la prueba de Stroop la función cognitiva del control inhibitorio está conservada.

Conclusiones

En este trabajo se describe un proceso para implementar un agrupador de señales EEG, el mismo que fue realizado mediante Python usando el algoritmo de *K-Means* de la librería *sklearn*.

Los resultados se analizan desde la perspectiva tanto técnica como psicológica y se pueden identificar ciertos patrones en los clusters con respecto a los resultados psicológicos. Los mejores resultados se obtienen con las señales en el dominio del tiempo, sin embargo, para lograr un mejor rendimiento del algoritmo con los datos en el dominio de la frecuencia, es necesario utilizar un mayor número de datos, ya que como se conoce, en este dominio se pueden extraer mayor cantidad de características de las señales, las mismas que pueden servir para detectar patrones no identificables con un set de datos pequeño.

Desde la perspectiva psicológica, se puede comprobar que el proceso de aprendizaje y de acomodación al momento de realizar una prueba psicológica, disminuye su potencial.

Finalmente, para trabajos futuros se plantea repetir el proceso con una cantidad de datos mayor, que permita por un lado validar lo realizado y, a su vez, mejorar los resultados obtenidos. De igual manera, es importante entrenar el algoritmo con datos de varios sujetos de diferentes edades, para nutrirlo con una mayor diversidad de muestras y señales, mejorando el rango del de análisis y su capacidad de obtención de conocimientos.

Agradecimientos

Los autores desean agradecer al Vicerrectorado de Investigaciones de la Universidad del Azuay por el apoyo financiero y académico, así como al personal de la escuela de Ingeniería de Sistemas y Telemática, y el Laboratorio de Investigación y Desarrollo en Informática (LIDI). Además, este trabajo es parte de los siguientes proyectos de investigación: “Fog Computing aplicado a monitoreo de dispositivos usados en ambientes de vida asistidos (Ambient Assisted Living). Caso de estudio: plataforma para el adulto mayor”; y “Diseño de arquitecturas y modelos de interacción para ambientes de vida asistida orientados a adultos mayores. Caso de estudio: ambientes lúdicos y sociales”. Ganadores de las convocatorias DIUC XVII y XVIII respectivamente, por ello los autores agradecen a la Dirección de Investigación de la Universidad de Cuenca (DIUC) por su apoyo.

Referencias

- Al-Nafjan, A., Hosny, M., Al-Ohali, Y., & Al-Wabil, A. (2017). Review and Classification of Emotion Recognition Based on EEG Brain-Computer Interface System Research: A Systematic Review. *Applied Sciences*, 7(12), 1239. <https://doi.org/10.3390/app7121239>
- Alqurashi, T., & Wang, W. (2019). Clustering ensemble method. *International Journal of Machine Learning and Cybernetics*, 10(6). <https://doi.org/10.1007/s13042-017-0756-7>
- Ayodele, T. O. (2010). Types of Machine Learning Algorithms. *New Advances in Machine Learning*.
- Basili, V. R., Caldiera, G., & Rombach, H. D. (1994). The goal question metric approach. *Encyclopedia of Software Engineering*, 2, 528-532. <https://doi.org/10.1.1.104.8626>
- Ben-Israel, D., Jacobs, W. B., Casha, S., Lang, S., Ryu, W. H. A., de Lotbiniere-Bassett, M., & Cadotte, D. W. (2020). The impact of machine learning on patient care: A systematic review. En *Artificial Intelligence in Medicine* (Vol. 103, p. 101785). Elsevier B.V. <https://doi.org/10.1016/j.artmed.2019.101785>
- Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers and Industrial Engineering*, 137, 106024. <https://doi.org/10.1016/j.cie.2019.106024>
- Emerenziani, G. Pietro, Vaccaro, M. G., Izzo, G., Greco, F., Rotundo, L., Lacava, R., Vignera, S. La, Calogero, A. E., Lenzi, A., & Aversa, A. (2020). Prediction equation for estimating cognitive function using physical fitness parameters in older adults. *PLoS ONE*, 15(5). <https://doi.org/10.1371/journal.pone.0232894>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD Process for Extracting Useful Knowledge from Volumes of Data. *Communications of the ACM*, 39(11), 27-34. https://dl.acm.org/doi/pdf/10.1145/240455.240464?casa_token=M2LKOWDtBmcAAAAA:Z_JvN1QHtQAdj4px-cISHRWzqwmFjrFq8cekNwgeEzMrYdMA7Rppimsq6Hxv5cwE_FqOMazrc4QPVw
- Likas, A., Vlassis, N., & J. Verbeek, J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, 36(2). [https://doi.org/10.1016/S0031-3203\(02\)00060-2](https://doi.org/10.1016/S0031-3203(02)00060-2)
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*. <https://doi.org/10.1037//0033-2909.109.2.163>
- Medina, B., Sierra, J. E., & Ulloa, A. B. (2018). Técnicas de extracción de características de señales EEG en la imaginación de movimiento para sistemas BCI Extraction techniques of EEG signals characteristics in motion imagination for BCI systems. *Espacios*, 39(22).
- Montañés, P., & de Brigard, F. (2011). *Neuropsicología clínica y cognoscitiva*. Centro Editorial, Facultad de Ciencias Humanas. https://books.google.com.ec/books?id=xW_DwAAQBAJ&printsec=frontcover&dq=neuropsicología+clínica&hl=es&sa=X&redir_esc=y#v=onepage&q=neuropsicología clínica&f=false
- Nasteski, V. (2017). An overview of the supervised machine learning methods. *HORIZONS.B*, 4. <https://doi.org/10.20544/horizons.b.04.1.17.p05>
- Onwuegbuzie, A., & Leech, N. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. En *International Journal of Social Research Methodology: Theory and Practice* (Vol. 8, Número 5, pp. 375-387). Routledge. <https://doi.org/10.1080/13645570500402447>
- Peña Barrientos, M. C., Gómez Mazo, T., Mejía Rúa, D., Hernández, J., & Tamayo Lopera, D. (2017). Caracterización del control inhibitorio en adolescentes del grado once de la Institución Educativa Normal Superior de Envigado-Colombia. *Psicoespacios*, 11(18). <https://doi.org/10.25057/21452776.906>
- Popov, T., Westner, B. U., Siltan, R. L., Sass, S. M., Spielberg, J. M., Rockstroh, B., Heller, W., & Miller, G. A. (2018). Time course of brain network reconfiguration supporting inhibitory control. *Journal of*

- Neuroscience*, 38(18). <https://doi.org/10.1523/JNEUROSCI.2639-17.2018>
- Priya, A., Garg, S., & Tigga, N. P. (2020). Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. *Procedia Computer Science*, 167. <https://doi.org/10.1016/j.procs.2020.03.442>
- Rennie, J. P., Zhang, M., Hawkins, E., Bathelt, J., & Astle, D. E. (2020). Mapping differential responses to cognitive training using machine learning. *Developmental Science*, 23(4). <https://doi.org/10.1111/desc.12868>
- Runeson, P., Höst, M., Rainer, A., & Regnell, B. (2012). Case Study Research in Software Engineering: Guidelines and Examples. En *Case Study Research in Software Engineering: Guidelines and Examples*. <https://doi.org/10.1002/9781118181034>
- Rusell, S., & Norvig, P. (2004). Inteligencia Artificial. Un Enfoque Moderno. En D. F. Aragón; (Ed.), *Inteligencia Artificial* (Segunda).
- Scarpina, F., & Tagini, S. (2017). The stroop color and word test. En *Frontiers in Psychology* (Vol. 8, Número APR). <https://doi.org/10.3389/fpsyg.2017.00557>
- Shani, R., Tal, S., Zilcha-Mano, S., & Okon-Singer, H. (2019). Can Machine Learning Approaches Lead Toward Personalized Cognitive Training? *Frontiers in Behavioral Neuroscience*, 13, 64. <https://doi.org/10.3389/fnbeh.2019.00064>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643-662. <https://doi.org/10.1037/h0054651>
- Stroop, J. R. (1992). Stroop 1933.Pdf. En *Journal of Experimental Psychology: General* (Vol. 121, pp. 15-23).
- Suryotrisongko, H., & Samopa, F. (2015). Evaluating OpenBCI Spiderclaw V1 Headwear's Electrodes Placements for Brain-Computer Interface (BCI) Motor Imagery Application. *Procedia Computer Science*, 72. <https://doi.org/10.1016/j.procs.2015.12.155>
- Trápaga, C., Pelayo, H., Sánchez, I., Bello, Z., & Bautista, A. (2018). *De la psicología cognitiva a la neuropsicología* (G. Moreno (ed.)). El Manual Moderno. https://books.google.com.ec/books?id=_T52zQEACAAJ&dq=psicología+cognitiva&hl=es&sa=X&redir_esc=y
- Wang, M., Zhang, S., Lv, Y., & Lu, H. (2018). Anxiety Level Detection Using BCI of Miner's Smart Helmet. *Mobile Networks and Applications*, 23(2). <https://doi.org/10.1007/s11036-017-0935-5>
- Xin, J. (2011). K-Means Clustering. *Encyclopedia of Machine Learning*, 563-564. https://doi.org/10.1007/978-0-387-30164-8_425
- S. Kaneta, I. Wakabayashi and T. Kawahara, "Feasibility of BMI improvement applying a Stroop effect," 2016 18th International Conference on Advanced Communication Technology (ICACT), 2016, pp. 681-684, doi: 10.1109/ICACT.2016.7423518.
- X. Hou, Y. Liu, O. Sourina, Y. R. E. Tan, L. Wang and W. Mueller-Wittig, "EEG Based Stress Monitoring," 2015 IEEE International Conference on Systems, Man, and Cybernetics, 2015, pp. 3110-3115, doi: 10.1109/SMC.2015.540.
- Lim, Alfred & Chia, Wai. (2015). Analysis of Single-Electrode EEG Rhythms Using MATLAB to Elicit Correlation with Cognitive Stress. *International Journal of Computer Theory and Engineering*. 7. 149-155. [10.7763/IJCTE.2015.V7.947](https://doi.org/10.7763/IJCTE.2015.V7.947).
- Elorriaga, Sergio. (2017). Estandarización Mexicana del Test Stroop: Datos Normativos.