

# Análisis del *clustering* para el estudio de registros encefalográficos en instancias de reposo precedentes y posteriores a una tarea de aprendizaje

Fabiana Alejandra Rossi

Magali Rodrigues Pires

Eliana Inés Padula

Álvaro López Malizia

## RESUMEN

La electroencefalografía (EEG) es una técnica no invasiva que detecta campos eléctricos a nivel de la corteza cerebral y permite encontrar patrones de activación cerebral. En particular, permite la identificación de patrones de activación anormales y diagnóstico precoz de determinadas patologías. El presente trabajo práctico tuvo como objetivo principal el estudio de la dinámica de los microestados para las condiciones previas y posteriores a una instancia de aprendizaje. A tal fin, se exploraron y calcularon los microestados para un conjunto de registros de EEG, y se determinó el número óptimo de agrupamientos que se necesitan para segmentar los datos mediante la técnica de k-means modificada. Se realizó una validación interna y externa de los clusters hallados utilizando diferentes metodologías (k-medoids y clustering jerárquico). No se encontró una coherencia elevada en los resultados arrojados por los distintos algoritmos. Finalmente, se profundizó sobre las transiciones que presentaron los microestados para cada instancia, y se encontró que la frecuencia de ocurrencia de dos estados particulares cambiaba al comparar las instancias de reposo en las condiciones pre y post aprendizaje. El procesamiento de los datos y confección de figuras se realizó utilizando el siguiente código: colab.

## I. INTRODUCCIÓN

La electroencefalografía (EEG) es una técnica que permite medir la actividad del cerebro mediante la detección de campos eléctricos a nivel de la corteza. Es un tipo de procedimiento no invasivo y registra de forma suavizada los potenciales extracelulares del tejido neuronal. Este registro espacio-temporal varía de forma organizada durante el tiempo (Lehmann et al., 1987). Durante períodos de descanso, se presentan patrones topográficos de actividad espontánea. Estos patrones discretos transicionan rápidamente entre sí reflejando la actividad neuronal global. La modificación de los patrones de actividad frente a distintos estímulos permite inferir la identidad del estímulo, y las características del proceso celular subyacente.

En este trabajo nos propusimos estudiar los registros electroencefalográficos correspondientes a períodos de reposo (período de "tarea negativa") previo y posterior a una instancia de aprendizaje.

Para ello, se visualizaron las series temporales para cada canal, se graficaron los promedios de la topografía para cada participante y se calcularon las componentes principales (PCA) para el total de los datos concatenados en el tiempo. Asimismo, se calculó el Campo de Potencia Global (GFP) y se registraron los máximos encontrados. Se realizó un Análisis de Componentes Principales con el objetivo de visualizar valores atípicos en los picos encontrados. Se calculó la distancia de Mahalanobis para cada uno de los registros en los picos de GFP hallados, y se realizó una prueba chi-cuadrado para determinar cuáles eran outliers de la distribución.

A continuación, y en virtud de los resultados del cálculo del parámetro Silhouette y el parámetro de Validación Cruzada (CV), se eligió el número óptimo de conjuntos (k) para agrupar los datos. Para dicho k, se realizó un perfil de Silhouette completo, y se calculó el porcentaje de los datos dentro de cada cluster que presentaban un valor negativo. Dicho perfil se comparó con aquel generado cuando no se consideraban los valores outliers definidos según su distancia de Mahalanobis.

Posteriormente, se aplicaron los métodos de clustering k-medoids y clustering jerárquico sobre los registros encefalográficos en los picos de GFP. En lo que refiere a las matrices de distancias, se utilizaron medidas que permitieron comparar la similitud angular entre los vectores y reflejar la correlación de las variables. Los prototipos generados en ambos casos fueron comparados con aquellos encontrados por el método de k-means modificado. La comparación entre los resultados obtenidos con cada método de clustering se realizó mediante el índice Rand.

Por último, se visualizaron los registros mediante la técnica de reducción de la dimensión TSNE (t-Distributed Stochastic Neighbor Embedding), se analizó el ajuste de los datos de EEG a los prototipos de microestados hallados previamente, y se calculó la frecuencia de ocurrencia de cada microestado prototípico en los registros encefalográficos analizados en las instancias de pre y post aprendizaje.

## II. DATOS

Se obtuvieron 10 registros encefalográficos de 5 sujetos. Hay dos archivos por participante, cada uno de ellos corresponde a al período de reposo previo al aprendizaje. Los registros se colectaron en 30 canales a una frecuencia de muestreo de 128 Hz.

## III. METODOLOGÍA

### A. Preprocesamiento

Los datos fueron preprocesados y se eliminaron las interferencias debido a los movimientos oculares y pestaños, mediante el uso del Análisis de Componentes Independientes (Delorme et al., 2007). Finalmente, los datos fueron filtrados por un filtro digital pasabanda entre 2 y 20Hz.

### B. Cálculo de GFP

El GFP (Global Field Power) es una medida de libre referencia que representa la fuerza del campo a nivel global. Los picos locales de la curva GFP representan los momentos de alto nivel global de neuro-sincronización. La escala se mantiene topográficamente estable alrededor de esas instancias. Esta medida se toma para encontrar el mapa topográfico del cluster dominante, y se calcula según la siguiente fórmula:

$$GFP = \sqrt{\frac{\sum_{i=1}^N (u_i - \bar{u})^2}{N}}$$

En primer lugar, se referenciaron los datos al promedio por cada muestra. Luego se calculó el GFP para cada uno de los registros encefalográficos, se identificaron los picos y se registraron los vectores del GFP para los máximos encontrados de todos los registros concatenados en el tiempo.

### C. Visualización e identificación de outliers

Los datos fueron visualizados mediante la técnica de componentes principales (PCA). Se consideraron las tres primeras componentes principales, ya que permiten explicar más del sesenta por ciento de la variabilidad observada. Para la identificación de outliers, se realizó el cálculo de la distancia de Mahalanobis. A partir de estos datos, se realizó el cálculo del p-valor de acuerdo a una distribución de chi-cuadrado. Los probabilidades inferiores a 0,001 se consideraron como puntos outliers.

### D. Cálculo de microestados: k-means modificado

Se utilizó el algoritmo k-means modificado para los registros EEG correspondientes a los picos de GFP. Se exploraron valores de cluster (k) comprendidos entre 2-10. Para garantizar la convergencia del algoritmo al agrupamiento óptimo se emplearon 1000 iteraciones.

Se empleó la métrica de Silhouette para cada uno de los agrupamientos generados de acuerdo al número de clusters seleccionados, y se calculó el parámetro de Validación Cruzada (CV). A partir de ambas métricas se determinó k=4 como número óptimo de clasificación. Adicionalmente, se confeccionó el gráfico de perfil completo de Silhouette para cada

uno de los clusters generados con k=4. Se determinó el porcentaje de valores cuyo coeficiente resultó negativo para esta métrica, incluyendo o no los valores atípicos determinados anteriormente.

### E. Validación externa

De forma complementaria, se aplicaron los algoritmos k-medoids y clustering jerárquico. El primero se seleccionó por tratarse de una opción robusta del k-means (al utilizar el cálculo de la mediana en vez de la media es menos sensible a la presencia de valores extremos). El clustering jerárquico es escogió por tratarse de un método aglomerativo (a diferencia del k-means que opera a partir de particiones).

La validación externa de los clusters determinados por k-means modificado se realizó a partir de la comparación de los agrupamientos generados por los distintos métodos de clustering mediante el índice de Rand ajustado.

### F. Representación de los registros en un dominio de menor dimensión.

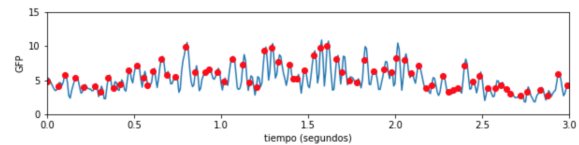
Los registros fueron visualizados mediante el método no lineal de T-SNE. Para la determinación efectiva de la reducción de la dimensión, y subsecuente representación bidimensional en las dos primeras componentes de T-SNE, se probaron distintos número de vecinos (parámetro "perplexity"). Se optó por utilizar el parámetro de complejidad igual a 50 para los datos con y sin outliers.

### G. Cálculo de frecuencia de microestados

Se calculó la frecuencia de ocurrencia de cada microestado prototípico en los registros encefalográficos analizados en los picos de GFP.

## IV. RESULTADOS

En primer lugar, se encontraron los picos máximos del GFP para los registros concatenados en el tiempo (Figura 1), y se guardaron los mapas para cada máximo. Se encontró un total de 13135 picos, distribuidos homogéneamente según indicado en la Tabla 1.

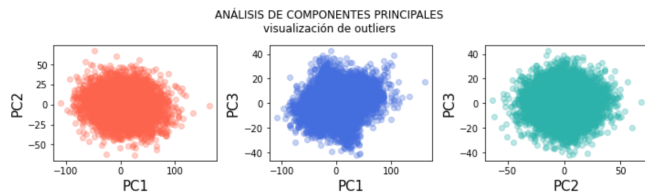


**Figura 1:** Fragmento de 3 segundos de los valores de GFP del sujeto número 13 en la condición post aprendizaje. En rojo se indican los máximos calculados.

Participante	pre-aprendizaje	post-aprendizaje
Sujeto 8	1293	1337
Sujeto 9	1366	1420
Sujeto 10	1289	1300
Sujeto 12	1176	1188
Sujeto 13	1353	1408

**Tabla 1:** Número de máximos GFP encontrados para cada uno de los sujetos experimentales en la condición pre y post aprendizaje.

Para cada máximo de GFP encontrado, se realizó un análisis de componentes principales (PCA), con el objeto de visualizar la presencia de outliers (Figura 2). Se graficaron las primeras tres componentes ya que explican aproximadamente el 67% de la variabilidad de los datos.

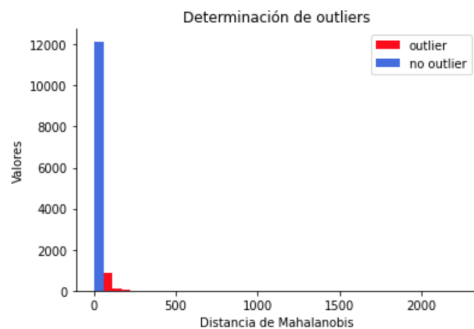


**Figura 2:** Resultado del análisis de los componentes principales aplicados sobre los registros encefalográficos coincidentes con los picos de GFP. De izquierda a derecha, se muestra la representación en dos dimensiones de las componentes principales 1-2 (anaranjado), 1-3 (azul) y 2-3 (verde).

Al analizar la Figura 2 se observa que existen algunos datos alejados de la distribución media de los mapas, correspondientes a aquellos valores que podrían ser considerados outliers por encontrarse hacia los extremos de los valores de cada componente analizada.

Dado que la determinación de valores atípicos mediante la técnica de PCA resulta inconveniente, se decidió calcular la distancia de Mahalanobis de los puntos al centro de la distribución, y considerar outliers a aquellos registros que presentaran un p-valor menor a 0.001 en la prueba de chi-cuadrado.

Como puede verse en la Figura 3, existen alrededor de 12000 registros que no son considerados outliers. Los registros que caen en dicha categoría representan el 7.41% de los datos analizados.

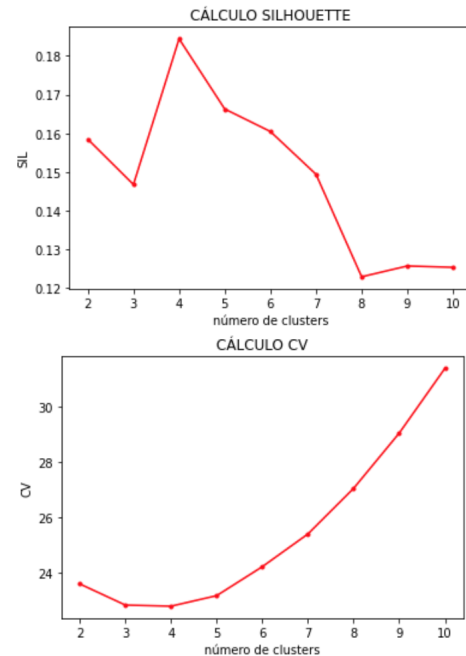


**Figura 3:** Cantidad de registros encefalográficos correspondientes a los picos de GFP según su distancia de Mahalanobis al centro de los datos. En rojo se indican los valores determinados como valores atípicos.

## 2. CALCULO DE MICROESTADOS UTILIZANDO K-MEANS MODIFICADO

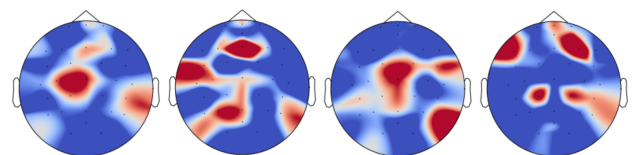
En un paso siguiente, se realizó el cálculo de microestados utilizando la técnica de k-means modificado visto en clase. Dado que la inicialización de esta metodología determina el

resultado final, se decidió realizar 1000 iteraciones de este algoritmo para asegurar una convergencia al valor óptimo. Se aplicó el método sobre los datos encefalográficos coincidentes con los picos de GFP, se utilizaron valores entre 2 y 10 para el número de clusters (k), y se calculó el parámetro de Silhouette y el coeficiente de validación cruzada para evaluar la metodología empleada (Figura 4).



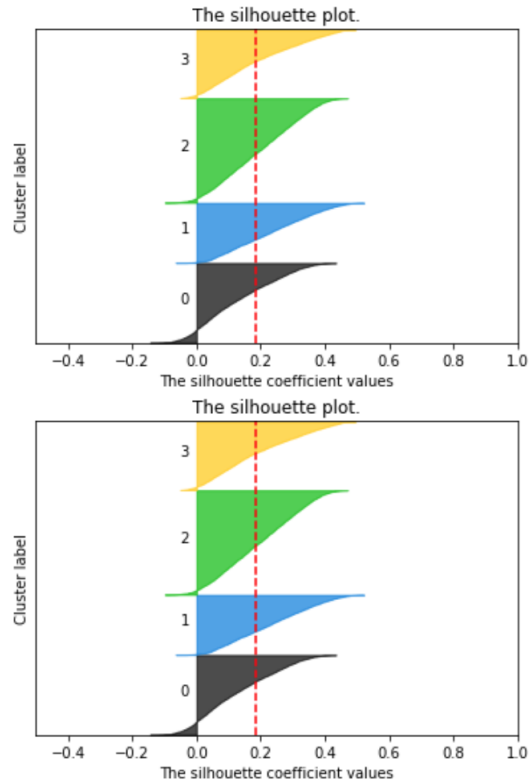
**Figura 4:** Cálculo del coeficiente de Silhouette (arriba) y de validación cruzada (abajo) para distintos valores de k (número de clusters) entre 2 y 10, sobre el mejor modelo resultante de 1000 iteraciones.

El coeficiente de Silhouette es una medida de validación interna que permite analizar el nivel cohesión alcanzado por los datos según el número de clusters. El parámetro de validación cruzada (CV) es una medida relacionada con la variación de los datos para cada agrupamiento. Al analizar la Figura 4, se observó que el mayor valor de Silhouette se encontraba para el agrupamiento de los datos en 4 clusters. Para este valor de k, el parámetro de validación cruzada presentó su menor valor. Por tal motivo, se utilizó este valor de k como óptimo. Las topografías características de los clusters generados se representan en la Figura 5.



**Figura 5:** Prototipos resultantes del agrupamiento de los mapas en 4 clusters. De izquierda a derecha se muestran los clusters "0" a "3".

Para el valor de  $k$  elegido, se realizó un perfil de Silhouette completo. De forma complementaria, se realizó el mismo análisis pero sin los valores atípicos (Figura 6). Un análisis más detallado de los valores del coeficiente Silhouette en cada cluster permitió calcular el porcentaje de datos con valor negativo para el análisis con los datos completos o sin los outliers (Tabla 2).



**Figura 6:** Perfil de coeficiente de Silhouette para los 4 clusters. Se grafican los datos completos (arriba) y los datos sin los outliers identificados por Mahalanobis (abajo). La línea punteada color rojo indica el coeficiente de Silhouette promedio en cada caso.

Cluster	datos completos	datos sin outliers
0	16.40	14.98
1	1.15	1.09
2	4.26	4.94
3	4.73	5.68

**Tabla 2:** Porcentaje de datos con valor de coeficiente de Silhouette negativo en cada cluster, al utilizar los datos completos incluyendo o no los outliers.

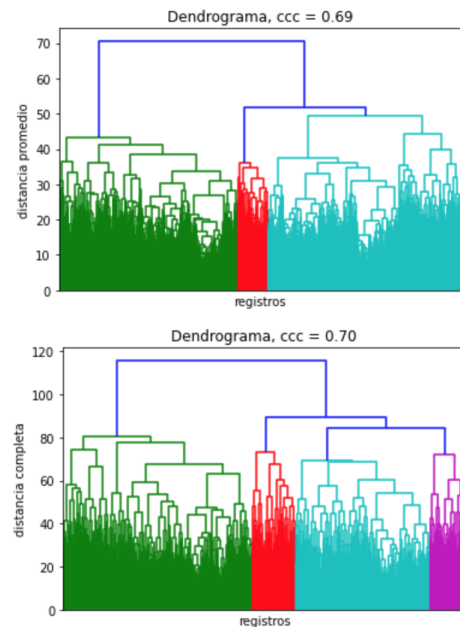
Tal como puede observarse en la Figura 6, la remoción del 7.41% de los datos correspondiente a los valores atípicos, no modificó sustancialmente el perfil de los coeficientes de Silhouette para cada cluster (Figura 6). El coeficiente promedio de Silhouette fue de 0.1834 y 0.1845 para los datos completos o sin outliers, respectivamente. En particular, para los clusters 0 y 1 se observó una disminución en el porcentaje de datos con valor de coeficiente de Silhouette negativo al eliminar los

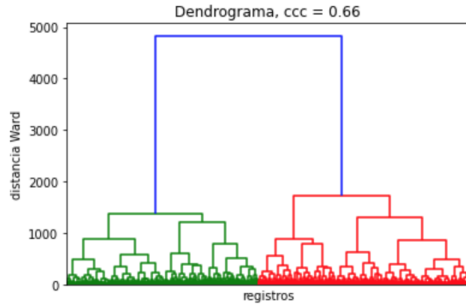
valores atípicos. Para el resto de los clusters, por el contrario, la eliminación de los valores atípicos produjo un aumento en el porcentaje de los valores negativos del coeficiente de Silhouette (Tabla 2). Si bien la remoción de los datos outliers mejoró levemente el coeficiente de Silhouette promedio, los mismos parecen indicar que los coeficientes negativos se deben, en mayor medida, a una inadecuada asignación de los datos a un dado cluster, y no tanto a la presencia de outliers. En este escenario, se debería aplicar un algoritmo de clusterización diferente, o bien repetir el método de  $k$ -means modificado sobre los datos sin outliers.

### 3. VALIDACIÓN EXTERNA DE LOS MICROESTADOS

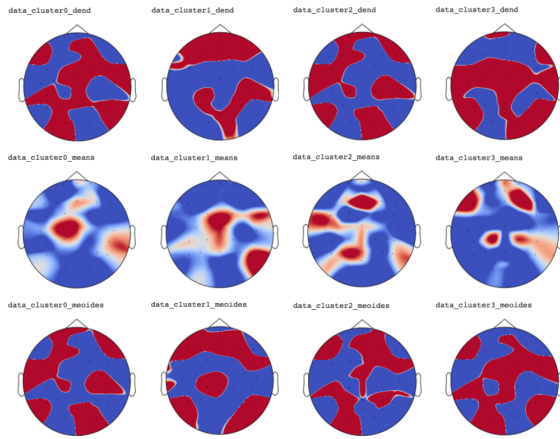
A continuación se aplicaron los algoritmos de clustering jerárquico y  $k$ -medoids utilizando los registros encefalográficos coincidentes con los picos de GFP. Dichas estrategias son métodos completos que asignan cada punto de manera exclusiva a un dado cluster. Se utilizó la matriz de distancia coseno entre los registros de la base de datos, dado que la misma permite inspeccionar correlaciones entre los perfiles EEG, para construir los dendrogramas jerárquicos. Se utilizaron distintas estrategias para confeccionar los dendrogramas (distancia promedio, completa o ward), y se generaron distinto número de clusters en cada caso (Figura 7). En particular, el dendrograma que utilizó la distancia completa para asignar puntos a cada cluster fue aquél con mayor coeficiente de correlación cofenético. Por tal motivo, se consideró el mismo para las siguientes comparaciones con el método de  $k$ -means modificado. Por otro lado, se aplicó la técnica de  $k$ -medoids sobre la matriz de distancias coseno calculada anteriormente. Se utilizó un  $k=4$  para generar los clusters.

Los prototipos promedio generados a partir de cada algoritmo se compararon con aquellos producidos por el método de  $k$ -means modificado (Figura 8).





**Figura 7:** Dendrogramas construidos a partir de la matriz de distancia coseno de los registros encefalográficos coincidentes con los picos de GFP, cuyas distancias entre clusters fueron la distancia promedio, completa o Ward.



**Figura 8:** Prototipos promedio generados a partir de la aplicación de los métodos de clustering jerárquico, k-means modificado o k-medoides.

Los resultados de la Figura 8 indican que ciertas topologías son coincidentes entre los métodos aplicados sobre los registros de la base de datos. En particular, los prototipos generados por los clusters utilizando los algoritmos de dendrograma y k-medoids muestran una gran similitud, aunque se presentan invertidos. El cluster 0, para sendos algoritmos presenta la mayor similitud topográfica. A fin de comparar los agrupamientos generados por cada modelo, se calculó el índice de Rand. El mismo se basa en un conteo de observaciones que son clasificados de la misma forma en particiones de distintos modelos. Los resultados se volcaron en la Tabla 3.

Cluster	Índice de Rand ajustado
k-medoides	0.105
clustering jerárquico	0.028

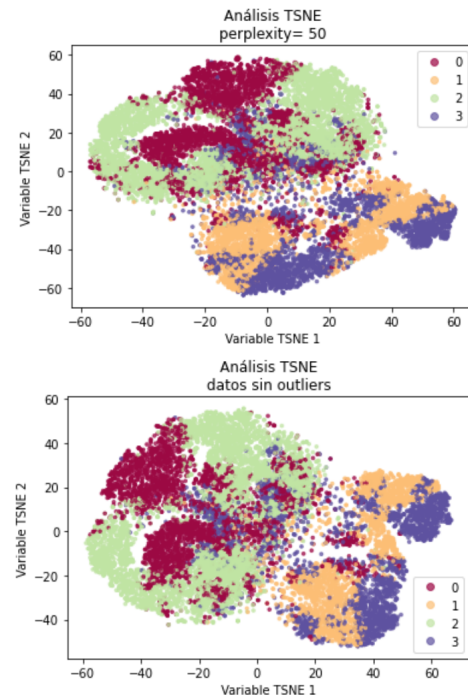
**Tabla 3:** Índice de Rand ajustado calculado para los pares de modelos k-means modificado versus k-medoids o dendrograma jerárquico (completo).

A partir de los datos de la Tabla 3, se puede concluir que el algoritmo de k-medoids es aquél que agrupa los registros de manera más parecida a la técnica de k-means modificada. No obstante, dado que el dominio del índice de Rand ajustado se encuentra entre 0 y 1, siendo 0 el valor que representa el menor

solapamiento de resultados y 1 el valor que representa un agrupamiento idéntico de los registros, puede concluirse que ambos métodos (k-medoids y clustering jerárquico) arrojan resultados que son pobremente coincidentes con los resultados generados al aplicar la técnica de k-means modificado.

#### 4. VISUALIZACIÓN

A fin de evaluar la posibilidad de un agrupamiento intrínseco de los datos, se utilizó la técnica de visualización de TSNE, que permite representar los datos en un menor número de dimensiones. Dicha técnica detecta relaciones no lineales entre las variables, y agrupa los datos considerando un cierto número de vecinos, en este caso, optimizado para 50. Los datos se colorearon de acuerdo al cluster correspondiente según el algoritmo de k-means modificado, y se comparó dicho resultado con el aparente agrupamiento generado al aplicar el algoritmo de TSNE sobre los datos sin los valores atípicos (Figura 9).



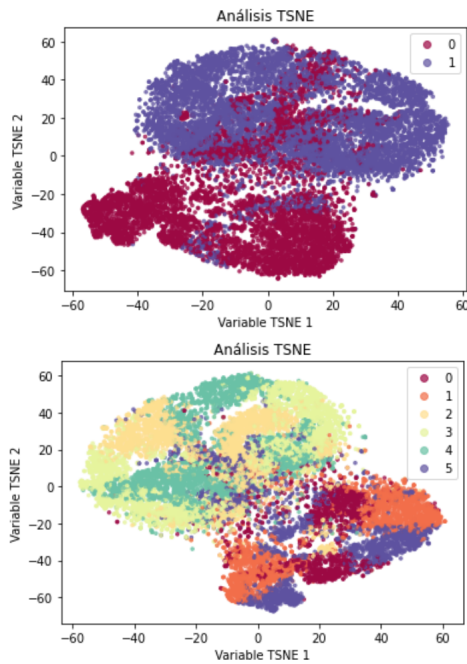
**Figura 9:** Representación de los registros EEG correspondientes a los picos de GFP en las dos primeras dimensiones del algoritmo TSNE. El color de cada punto representa el cluster original asignado por el algoritmo de k-means modificado. Arriba se utilizaron datos completos, abajo se muestran los resultados utilizando datos sin los outliers.

Tal como puede observarse en la Figura 9, no parece haber un agrupamiento de los datos que se solape con los clusters obtenidos utilizando el algoritmo de k-means modificado.

Con el objeto de realizar una comparación más completa entre los agrupamientos generados por ambos métodos, se aplicó la técnica de TSNE sobre los datos completos, y se colorearon los puntos de acuerdo al cluster asignado por el



método de k-means modificado cuando k era igual a 2 o 6 (Figura 10).



**Figura 10:** Representación de los registros EEG correspondientes a los picos de GFP en las dos primeras dimensiones del algoritmo TSNE. El color de cada punto representa el cluster original asignado por el algoritmo de k-means modificado. Arriba se presenta el gráfico correspondiente a  $k=2$ , mientras que abajo se presenta el gráfico con  $k=6$

Los resultados de las Figuras 9 y 10 indican que existe una mayor coincidencia entre el agrupamiento aparente de los datos (visualizado por TSNE) y los clusters generados por k-means modificado cuando  $k=2$ , ya que los dos centros de datos más densos observados en la representación de las dos primeras componentes del TSNE están coloreados principalmente por uno de los dos colores, que indican el cluster de k-means.

## 5. CÁLCULO DE FRECUENCIA DE MICROESTADOS

Por último, se decidió calcular la frecuencia de los microestados identificados por la técnica de k-means modificado en los registros EEG coincidentes con los picos de GFP, en las condiciones pre y post aprendizaje, por separado (Tabla 4).

Cluster k-means	pre-aprendizaje	post-aprendizaje
0	0.254	0.255
1	0.334	0.333
2	0.183	0.199
3	0.227	0.211

**Tabla 4:** Frecuencia de registros EEG en los picos de GFP asignados a los cluster 0 a 3 de acuerdo al algoritmo k-means modificado, en la condición pre y post aprendizaje.

Tal como puede observarse en la Tabla 4, la frecuencia de aparición de los clusters 0 y 1 en los registros pre y post aprendizaje son semejantes. Por el contrario, la frecuencia de los clusters 2 y 3 varía sustancialmente de acuerdo a la instancia de reposo: se produce un aumento en la frecuencia del cluster 2 y una disminución de aquella correspondiente al cluster 3 en los datos del período de reposo post aprendizaje (respecto el período pre-aprendizaje).

## CONCLUSIONES

El presente trabajo tuvo como objetivo principal el estudio y exploración de datos de EEG registrados durante períodos de reposo previo y posterior a un aprendizaje. Para ello, en primer lugar, se calculó el valor de GFP para los registros concatenados en el tiempo. A partir de estos valores, se calcularon los picos correspondientes, sobre los cuales se aplicaron distintos algoritmos de clustering.

La visualización de los datos en un menor número de dimensiones determinadas por PCA, permitió advertir la presencia de un pequeño porcentaje de valores outliers. Se utilizó la distancia de Mahalanobis para identificar los mismos. A continuación, se aplicó la técnica de k-means modificado sobre los datos completos. Se utilizó la correlación de los datos como una medida de distancia entre ellos, y se calculó el coeficiente de Silhouette junto con el parámetro de validación cruzada (CV) para determinar el número óptimo de conjuntos que se necesitan para segmentar los datos (validación interna).

A partir de estas métricas, se eligió el número de clusters igual a 4, y se estudió el perfil de Silhouette completo incluyendo o no los valores identificados como outliers. Se observó que no existían diferencias marcadas al remover los outliers (leve aumento del coeficiente de Silhouette promedio), lo que permitió suponer que valores negativos de Silhouette podrían deberse a una inadecuada asignación de los valores a cada cluster.

A continuación, se realizó una validación externa de los clusters originados por k-means modificado mediante la comparación con los agrupamientos generados por otros dos métodos. Este tipo de revisión de los clusters otorga robustez sobre la clasificación cuando se obtienen conclusiones similares por métodos diferenciales.

En primer lugar, se calculó la matriz de distancias coseno, que permite identificar el grado de correlación de los registros de nuestra base de datos, y se utilizó la misma para generar clusters mediante el algoritmo de k-medoids o clustering jerárquico. En el primer caso, se utilizó un número de clusters igual a 4, mientras que en el caso del agrupamiento jerárquico, se utilizaron distintas estrategias para calcular las distancias entre los clusters generados: distancia promedio, completa o Ward. Dado que el coeficiente cofenético de los dendrogramas generados resultó mayor para la opción de distancia completa, se decidió utilizar esta última noción de distancia.

La comparación de los resultados generados se realizó de a pares utilizando el índice de Rand ajustado. El mismo se basa en un conteo de observaciones que son clasificados coincidentemente en particiones pertenecientes a distintos modelos. Con

relación al modelo de k-means modificado, ambos métodos presentaban un índice de Rand cercano a 0. Esto indicó que los métodos agrupan los registros de manera distinta al clustering generado por k-means modificado. Este resultado, sugiere que los grupos generados por k-means, no presentarían una identidad real robusta. Una alternativa que podría ser evaluada es la repetición del análisis, ahora mediante un  $k=2$ , que presenta un máximo local dentro del análisis del Silhouette (Figura 4).

Continuando con la caracterización de los grupos, se decidió aplicar un método de visualización no supervisado que detecta relaciones no lineales: TSNE. Dicho algoritmo, reforzó la idea de un agrupamiento intrínseco de los datos diferente a los clusters originados por k-means modificado (cuando se utiliza el  $k$  óptimo igual a 4). De hecho, los resultados indicaron que los datos aparentan agruparse en dos conjuntos, y que los mismos son mayormente coincidentes con las etiquetas generadas cuando la elección del valor  $k$  es igual a dos. Esta evidencia gráfica, en consonancia con la validación externa, apoya nuevamente la idea de repetir la validación interna ahora con una cantidad de clusters igual a 2.

Por último, se decidió evaluar la dinámica de los microestados para las condiciones previa y posterior al aprendizaje. Para ello, se calculó la frecuencia de los registros EEG correspondientes con los picos GFP que fueron etiquetados con distintos clusters según k-means. Se observó que para los clusters 0 y 1 no existía una diferencia marcada entre los períodos de reposo pre y post aprendizaje. Por el contrario, se observó un aumento en la frecuencia de asignación del cluster 2 en la condición post aprendizaje, y una disminución en la frecuencia de asignación del cluster 3, respecto la condición pre-aprendizaje. De acuerdo a nuestros resultados, la cantidad de microestados que permitirían modelar las actividad neuronal bajo estas condiciones serían 2. Llamativamente, los elementos de ambos clusters se distribuyen en regiones opuestas dentro del gráfico de la Figura 9.

En conclusión, la aplicación de métodos de agrupamiento sobre series temporales de registros encefalográficos permite identificar estados de actividad particulares del cerebro en determinadas condiciones experimentales. En sintonía con la teoría de la existencia finita de microestados que permiten modelar la actividad cerebral, nuestro análisis permitió detectar dos estados que condensarían la variabilidad detectada entre las instancias de pre y post aprendizaje. A pesar de que nuestras clasificaciones carecen de robustez, la visualización de la información por el método de TSNE, indica la presencia de una cantidad de estructuras finitas que se obtienen a partir de los datos encefalográficos y acompañan la teoría general.

#### REFERENCIAS

1- D. Lehmann, H. Ozaki, I. Pal, EEG alpha map series: brain micro-states by space-oriented adaptive segmentation, *Electroencephalography and Clinical Neurophysiology*, Volume 67, Issue 3, 1987, Pages 271-288, ISSN 0013-4694, [https://doi.org/10.1016/0013-4694\(87\)90025-3](https://doi.org/10.1016/0013-4694(87)90025-3).

2- Delorme A, Sejnowski T, Makeig S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage*. 2007 Feb 15;34(4):1443-9. doi: 10.1016/j.neuroimage.2006.11.004. Epub 2006 Dec 26. PMID: 17188898; PMCID: PMC2895624.