



INSTITUTO DE ESTUDIOS SUPERIORES DE MONTERREY
CAMPUS GUADALAJARA

Análisis de datos de actividad cerebral

Por:

Juan Pablo Bernal Lafarga – A01742342

TC2034.301 – Modelación del aprendizaje con inteligencia artificial (Gpo 301)

Dr. Omar Mendoza Montoya

Junio 12, 2023

Zapopan, Jalisco

1. Una descripción detallada de los datos obtenidos en los dos experimentos.

Una electroencefalografía o EEG es una prueba simple e indolora que registra la actividad eléctrica del cerebro para comprobar el buen funcionamiento de las distintas áreas del mismo. Estos EEG se emplean para detectar problemas, eso es su uso médico, pero también son usadas para encontrar parámetros o lecturas recurrentes de diferentes áreas del cerebro a ciertos estímulos para poder crear una asociación causa-efecto entre las reacciones del cerebro y los eventos exteriores.

El intercambio de información entre las células del sistema nervioso se da por medio de corrientes eléctricas. En conjunto, esta actividad neural da origen a variaciones del potencial eléctrico en el espacio extracelular, variaciones en espacio y tiempo que pueden ser registradas y estudiadas para caracterizar sistemas neuronales. Estas oscilaciones en el potencial eléctrico local se han categorizado en diferentes bandas de frecuencia bien definida o ritmos que van aproximadamente desde 0.05 Hz hasta 500 Hz. Conforme aumenta el ritmo, podemos notar el nivel de actividad del cerebro.

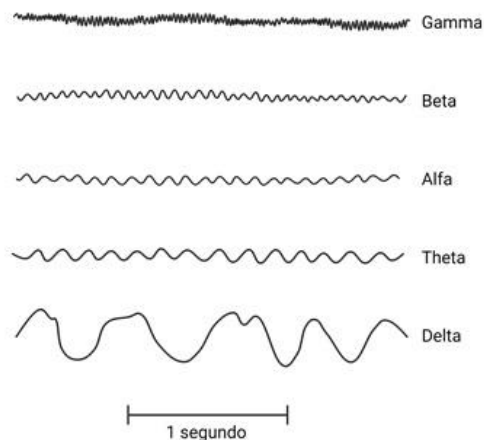
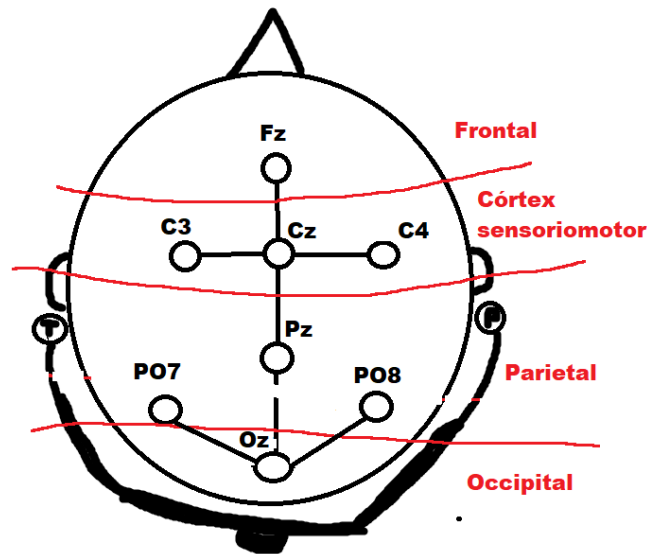


Figura 1. Ritmos del electroencefalograma.

El equipo que nosotros usamos es el Unicorn Hybrid Black, una gorra EEG portátil de 8 canales para interfaz cerebro-computadora para adquirir perfectamente las ondas cerebrales. Realiza muestreo con 24 bits y 250 Hz por canal.

La posición de los electrodos se divide en 4 secciones.

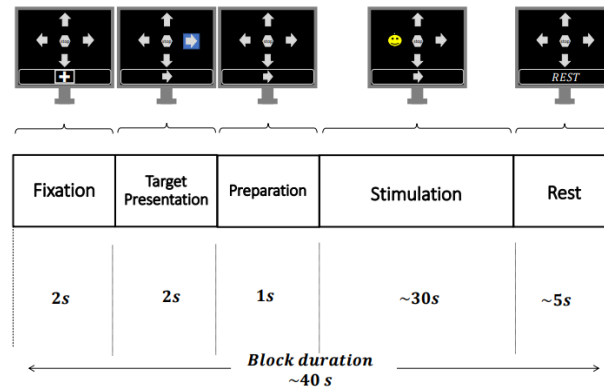
1. Córtex sensoriomotor para realizar BCIs basados en imágenes motoras
2. Sobre las áreas central, parietal y occipital para los paradigmas P300
3. Sobre las regiones parietales para los paradigmas de potencial visual evocado de estado estable y VEP basados en códigos
4. Los electrodos de potencial y tierra se fijan en los mastoides.



El rango de señales típico es de $-100\mu V$ a $100\mu V$, todo lo que nos dé fuera de tal rango es ruido.

Llevamos a cabo dos experimentos a los cuáles llamamos “P300” y “Tareas Cognitivas”. El P300 es un experimento que se da de observar el número de veces que una cruceta direccional coloca de manera random una carita feliz en una flecha que el sistema nos pide observar. La interfaz BCI capta las señales cerebrales y las interpreta dependiendo cómo se comportan nuestras señales o ritmos a la hora de mirar hacia un lado y medir nuestra reacción al hacer un conteo.

El proceso de extracción de características que sigue el P300 es: Hace aparición la carita feliz y hay un corte de 800 milisegundos, en el cual habrá 200 muestras por canal. Luego, se aplica un filtro para eliminar bandas de frecuencia para separar y analizar. Este filtro es de 4 Hz a 20 Hz. Después, se aplica un submuestreo para balancear resultados y nos quedamos con 50 muestras por canal. Finalmente, se aplica un filtrado espacial.



El experimento de “Tareas Cognitivas” consistió en 12 ejercicios diferentes, los cuales podemos clasificar en 3 tipos: tareas de cálculo, tareas de lectura y tareas de nombres de cosas. El propósito de este experimento era lograr clasificar nuevas observaciones de forma efectiva, es decir, que la computadora identificara las señales o ritmo del cerebro a la hora de resolver una tarea de cálculo, lectura o de nombres (podemos asociarlo a la memoria).

El experimento de "Tareas Cognitivas" se maneja de la siguiente manera: Hay tres fases las cuales son no realizar ninguna tarea, realizar una tarea y descansar, Esto se repite de manera secuencial hasta que se completan las 12 tareas existentes. Cada espacio de no realizar y realizar es de 10 cada uno, y el descanso es de 5. De estas 3 fases, extraemos los datos de realizar y no realizar tareas, para así obtener 13 clases diferentes.

Luego, en todos los canales se realiza un análisis de potencia espectral, el cual analiza el nivel de energía aportado por cada frecuencia. Finalmente, se utiliza programación para organizar el vector de características para el clasificador.

Resuelve mentalmente

$21 \div 3 = ?$	$54 \div 6 = ?$
$45 \div 5 = ?$	$48 \div 2 = ?$
$39 \div 3 = ?$	$55 \div 5 = ?$
$77 \div 7 = ?$	$42 \div 7 = ?$
$36 \div 2 = ?$	$48 \div 4 = ?$

Lee en silencio el siguiente texto:

El rayo de la muerte de Tesla era un arma teórica que permitía disparar un haz de partículas microscópicas hacia seres vivos u objetos para destruirlos. Supuestamente fue inventado entre la década de 1920 y 1930 por Nikola Tesla, un científico e inventor serbio que emigró a los Estados Unidos. Tesla decía haber inventado un rayo de la muerte usando lo que él denominaba "teleforce", una forma de transmitir energía eléctrica sin cables. El aparato nunca fue desarrollado ni demostrado públicamente, pero ha alimentado la imaginación de muchos autores de ciencia ficción y ha inspirado la creación de conceptos como la pistola de rayos láser.

Imagina nombres de personas que comiencen con la letra M

Para la extracción de características, la Densidad Espectral de Potencia (PSD) de una señal es una función matemática que da a conocer la distribución de la potencia de dicha señal sobre las distintas frecuencias en donde está formada.

Al final, del experimento “P300” obtuvimos dos clases que son Potencial P300 vs No Potencial P300, es decir, si estamos contando la cuántas veces aparece la carita feliz en la flecha indicada, o si no estamos haciendo nada (que en el experimento es una cruz). Y en el experimento “Tareas Cognitivas” obtuvimos 13 clases, de las cuales las primeras 4 representan las tareas de nombres de cosas, de la 5-8 son tareas de lectura y de 9-12 son tareas de cálculo, mientras que la 13va tarea es nuevamente no hacer nada (representada por una cruz).

2. Un resumen ejecutivo de los resultados obtenidos en la evaluación de los clasificadores probados, incluyendo la obtención de hiperparámetros y selección de características.

Podremos observar que el clasificador que más se adapta a los datos en todos nuestros casos es el Gradient Boosting gracias a su capacidad para manejar características complejas, el que haga una reducción de ruido y mejora de la precisión de los datos del EEG, y su gestión de balanceo de clases, pues además de presentarse como el clasificador con más accuracy en los 3 casos, también es el que mejor balanceo de recall presenta en sus clases.

	precision	recall	f1-score	support
1	0.75	0.72	0.73	280
2	0.73	0.75	0.74	280
accuracy			0.74	560
macro avg	0.74	0.74	0.74	560
weighted avg	0.74	0.74	0.74	560

	precision	recall	f1-score	support
1	0.70	0.57	0.63	56
2	0.70	0.75	0.72	56
3	0.58	0.64	0.61	56
accuracy			0.65	168
macro avg	0.66	0.65	0.65	168
weighted avg	0.66	0.65	0.65	168

En general, los resultados obtenidos en el caso número 3 (tareas de cálculo vs tareas de lectura vs tareas de nombres de cosas) no fueron tan buenos, esto podría deberse a la complejidad de un clasificador cuando cuenta con más de 2 clases en los datos proporcionados, por lo que estaríamos entrando en el mundo de los clasificadores multiclases y sus problemas

tales como el desbalanceo del recall de las clases, la interferencia entre las mismas clases y la dificultad de interpretación de resultados.

Además del Gradiente Boosting, los demás clasificadores mostraron buenos resultados en los primeros 2 casos (P300 y no tarea vs tarea), pero al momento del 3er caso, todos estos presentaban resultados muy bajos. De entre 39-50 de accuracy aproximadamente. Pero nuevamente, los datos de un clasificador multiclase son difíciles de interpretar.

El método utilizado para obtener los hiperparámetros óptimos fue el GridSearchCV, una técnica exhaustiva que prueba todas las combinaciones posibles de valores de hiperparámetros de un modelo. En nuestro caso, el modelo con el que usamos esta técnica fue KNN, y los hiperparámetros a optimizar fueron el número de posibles vecinos (n_neighbours), los valores posibles para el peso de los vecinos (weights) y las posibles distancias utilizadas (p), ya sea Manhattan o Euclidiana. Hubo una mejoría en la exactitud de los modelos KNN de cada caso presentado.

	precision	recall	f1-score	support
1	0.61	0.64	0.63	280
2	0.62	0.59	0.61	280
accuracy			0.62	560
macro avg	0.62	0.62	0.62	560
weighted avg	0.62	0.62	0.62	560

```
Mejores hiperparámetros encontrados:  
{'n_neighbors': 7, 'p': 2, 'weights': 'uniform'}  
Exactitud del modelo con los mejores hiperparámetros: 78.57%
```

Finalmente, para selección de características utilizamos el método de filtrado, el cual evalúa la relevancia de cada característica de forma individual. Luego, estas características se ordenan de acuerdo con algún índice de relevancia, de forma que las características de valor más bajo pueden ser eliminadas. Para el caso de clasificación, el método para determinar la relevancia de las características es `f_classif`, el cual estima el grado de dependencia lineal entre dos variables, de manera que, si la relación entre la característica evaluada y el target es no lineal, no es considerada.

Las características que esperábamos del método de filtrado eran 10. Comparamos el accuracy de cada clasificador tomando todas las características, y con las características filtradas. Como se observará a continuación, obtuvimos buenos resultados.

	precision	recall	f1-score	support
1	0.87	0.81	0.84	280
2	0.82	0.88	0.85	280
accuracy			0.84	560
macro avg	0.84	0.84	0.84	560
weighted avg	0.84	0.84	0.84	560

Este es el resultado del clasificador SVM con los datos del experimento “P300.

	precision	recall	f1-score	support
1	0.84	0.80	0.82	280
2	0.81	0.85	0.83	280
accuracy			0.82	560
macro avg	0.82	0.82	0.82	560
weighted avg	0.82	0.82	0.82	560

Y este es el resultado del clasificador con las características consideradas más relevantes gracias al método de filtrado de características. Como podemos observar, usando solo 10 de las 151 características, seguimos obteniendo un accuracy de 0.82 y un buen balanceo de las clases.

3. Seleccione uno de los clasificadores evaluados que no se haya visto en clase, y describa la teoría de su funcionamiento en no más de dos páginas (mínimo una).

Un modelo de potenciación del gradiente (Gradient Boosting) está formado por un conjunto de árboles de decisión individuales, entrenados de forma secuencial, de manera que cada nuevo árbol trata de mejorar los errores de los árboles anteriores. La predicción de una nueva observación se obtiene agregando las predicciones de todos los árboles individuales que forman el modelo.

Si bien es cierto que el proceso de boosting del Gradiente Boosting consigue mejorar la capacidad predictiva en comparación a los modelos basados en un único árbol de decisión, esto tiene un costo asociado, la interpretabilidad del modelo se reduce. Al tratarse de una combinación de árboles, no es posible obtener una representación gráfica sencilla del modelo y no es inmediato identificar de forma visual que predictores son más importantes.

Gradient Boosting es una generalización del algoritmo AdaBoost que permite emplear cualquier función de costo, siempre que sea diferenciable. La flexibilidad de este algoritmo ha hecho posible aplicar boosting a problemas de regresión, clasificación múltiple, entre otros. Convirtiéndolo en uno de los métodos de ML de mayor éxito.

Se ajusta un primer weak learner f_1 con el que se predice la variable y , y se calculan los residuos $y - f_1(x)$. A continuación, se ajusta un nuevo modelo f_2 , que intenta predecir los residuos del modelo anterior, es decir, trata de corregir los errores que ha hecho el modelo f_1 .

$$f_1(x) \approx y$$

$$f_2(x) \approx y - f_1(x)$$

En la siguiente iteración, se calculan los residuos de los dos modelos de forma conjunta $y - f_1(x) - f_2(x)$, los errores cometidos por f_1 y que f_2 no ha sido capaz de corregir, y se ajusta un tercer modelo f_3 para tratar de corregirlos. Este proceso se repite M veces, de forma que cada nuevo modelo minimiza los errores del anterior.

$$f_3(x) \approx y - f_1(x) - f_2(x)$$

Dado que el objetivo del Gradient Boosting es ir minimizando los errores iteración con iteración, es susceptible a overfitting. Pero, una forma de evitar este problema es utilizando un valor de regularización que se conoce como learning rate (λ), que limita la influencia de cada los modelos en el conjunto del ensemble. Como consecuencia de esta regularización, se necesitan más modelos para formar el ensemble, pero se consiguen mejores resultados.

$$f_1(x) \approx y$$

$$f_2(x) \approx y - \lambda f_1(x)$$

$$f_3(x) \approx y - \lambda f_1(x) - \lambda f_2(x)$$

$$y \approx \lambda f_1(x) + \lambda f_2(x) + \lambda f_3(x) + \dots + \lambda f_m(x)$$

Los hiperparámetros más importantes del Gradiente Boosting son el número de árboles, esto debido a que si aumentamos el número de árboles sin razón alguna corremos el riesgo de hacer overfitting, y el learning rate es el otro hiperparámetro crítico, pues es el que permite cómo de rápido aprende el modelo y si lo aumentamos, corremos nuevamente el riesgo de tener

overfitting. Estos dos hiperparámetros son interdependientes, cuánto menor es el learning rate, mayor número de árboles necesitamos para obtener buenos resultados, pero reducimos el riesgo de tener overfitting.

4. Una conclusión de este trabajo, junto con tu respuesta a las siguientes preguntas:

- a. ¿Consideras que el monitoreo de la actividad cerebral de una persona es benéfico para su salud?

El uso de los EEG en la medicina es vital para detectar posibles problemas en el cerebro, además de que puede advertirnos acerca de alguna falla en otra parte del cuerpo, como mostrar anomalías en la región del cerebro asociada a la actividad motriz.

- b. ¿Qué otro tipo de aplicaciones basadas en el monitoreo de datos médicos podrían ser benéficos para la salud?

No solo se limita a detectar enfermedades, sino que también ayuda en el difícil caso de la asociación de actividades o movimientos a ciertos ritmos o actividades cerebrales, pudiéndose estas codificar e implementar en sistemas que pueden funcionar como prótesis.

- c. ¿Qué otro tipo de información, variables o señales se podrían monitorear con estos dispositivos para desarrollar aplicaciones médicas?

Bueno, como ya observamos, además de las señales cerebrales presentadas por el EEG, también se pueden registrar señales fisiológicas simultáneamente, como el electrocardiograma, la electromiografía, la electrooculografía, entre otros. Para que, al integrar todas estas señales, se pueda proporcionar una visión más completa de lo que está ocurriendo dentro del paciente y permitir un análisis más profundo.

Un caso en donde podemos ver el uso de estas señales en conjunto es en las polisomnografías. Un estudio del sueño que registra ciertas funciones corporales a medida que uno duerme o trata de dormir. Lo que se registra en el estudio es el flujo de aire en los pulmones, los niveles de oxígeno en la sangre, la posición del cuerpo, las ondas cerebrales (EEG), el esfuerzo y frecuencia respiratoria, la actividad eléctrica de los músculos (EMG), los movimientos oculares (EOG) y la frecuencia cardíaca (ECG).

Se utiliza para diagnosticar trastornos del sueño.

CONCLUSIÓN:

A manera de conclusión, La lectura de señales fisiológicas y la extracción de características de estas desempeñan un papel muy importante en el ámbito de la salud y la investigación. Estas señales proporcionan información valiosa sobre el funcionamiento del cuerpo y su estado, lo que permite la comprensión más profunda de la salud y el bienestar de las personas. Son valiosas tanto para el diagnóstico y detección de enfermedades, monitoreo de la salud y prevención de la salud, como para la personalización de tratamientos y la búsqueda de codificar el cerebro humano y cómo funciona, de tal manera que las prótesis y las máquinas puedan alcanzar un nuevo nivel de sofisticación y utilidad.

También, los clasificadores son esenciales para aprovechar al máximo la utilidad de las señales fisiológicas en el ámbito de la salud. Estos algoritmos de aprendizaje automático permiten el análisis y la interpretación de los datos de las señales, lo que facilita el diagnóstico de enfermedades, el monitoreo de la salud, la personalización del tratamiento y la predicción de resultados. Su integración en la práctica médica puede mejorar la precisión de los diagnósticos, optimizar los tratamientos y mejorar los resultados de salud de los pacientes.

Aprendizaje por transferencia

Para esta etapa del proyecto, consideren únicamente los datos del experimento de P300. Para al menos dos clasificadores probados en el paso anterior, entrenen el modelo con uno de los archivos de datos de uno de los sujetos, y prueben dicho modelo con otro de los archivos de datos de otra persona. Reporte el rendimiento del modelo.

Conteste las siguientes preguntas:

1. ¿Es posible entrenar un modelo de clasificación con los datos de un sujeto y obtener buenos resultados con los datos de otro sujeto? ¿Por qué?

Es posible obtener buenos resultados, pero depende de factores como la similitud de los sujetos, la variabilidad de los datos, pues se registran diferentes niveles de actividad cerebral en cada sujeto debido a la concentración de cada uno, también se debe de considerar las características importantes del archivo de cada sujeto. Si observamos el filtrado para selección de características de cada sujeto para el experimento P300, podemos notar que las características elegidas no son las mismas para ambos casos, por lo que podría haber un gran decremento del accuracy a la hora probar los datos del sujeto 2 en el modelo hecho con los datos del sujeto 1.

```
Resultados de la validación cruzada en los datos del primer sujeto:  
Precisión promedio: 0.87219514710812  
Precisión en los datos del segundo sujeto: 0.5660714285714286
```

2. ¿Consideras que debes hacer ajustes cuando tengas datos de otro sujeto?

Consideraría hacer ajustes en los hiperparámetros del modelo, podría probarlos al azar, pero podemos utilizar métodos como el GridSearchCV para encontrar los hiperparámetros óptimos para el modelo. Y también ajustaría la selección de características, pues como mencioné anteriormente, ambos archivos del experimento P300 de los dos sujetos muestran características relevantes diferentes, por lo que, si el modelo hace énfasis en utilizar las mismas características que con el sujeto 1, el rendimiento y resultados del archivo del sujeto 2 se verá afectado en el intento.

Referencias:

- F. Villalba (fervilber@gmail.com). (n.d.). *Capítulo 5 Naive Bayes- clasificación bayesiano ingenuo / Aprendizaje supervisado en R*. <https://fervilber.github.io/Aprendizaje-supervisado-en-R/ingenuo.html>
- Gradient Boosting con python*. (n.d.).
https://www.cienciadedatos.net/documentos/py09_gradient_boosting_python
- Saini, A. (2023). Gradient Boosting Algorithm: A Complete Guide for Beginners. *Analytics Vidhya*. <https://www.analyticsvidhya.com/blog/2021/09/gradient-boosting-algorithm-a-complete-guide-for-beginners/#:~:text=What%20is%20a%20Gradient%20boosting,errors%20of%20the%20previous%20model>.
- Manuales MSD. (2023, June 5). *Datos clave: Electromiografía (EMG) y estudios de conducción nerviosa - Manual MSD versión para público general*. Manual MSD Versión Para Público General. <https://www.msdmanuals.com/es-mx/hogar/breve-informaci%C3%B3n-trastornos-cerebrales,-medulares-y-nerviosos/diagn%C3%B3stico-de-los-trastornos-cerebrales-medulares-y-nerviosos/electromiograf%C3%ADa-emg-y-estudios-de-conducci%C3%B3n-nerviosa>
- Fisiología de la actividad eléctrica del cerebro / FISIOLOGÍA*. (n.d.).
<https://fisiologia.facmed.unam.mx/index.php/fisiologia-de-la-actividad-electrica-del-cerebro/>
- Breinbauer, S. (2021). Tecnología | unicorn-bi.com. *unicorn-bi.com*. <https://www.unicorn-bi.com/es/brain-interface-technology/>
- Luengas, L. A., & Toloza, D. C. (2020). Análisis frecuencial y de la densidad espectral de potencia de la estabilidad de sujetos amputados. *Tecno Lógicas*, 23(48), 1–16.
<https://doi.org/10.22430/22565337.1453>
- KeepCoding, R. (2022, December 9). Métodos de filtrado para selección de características en machine learning. *KeepCoding Bootcamps*. <https://keepcoding.io/blog/metodos-de-filtrado-para-seleccion-de-caracteristicas/>
- Polisomnografía: MedlinePlus enciclopedia médica*. (n.d.).
<https://medlineplus.gov/spanish/ency/article/003932.htm>