



**UNIVERSIDAD
DE ANTIOQUIA**

**Técnicas de filtrado espacial de señales EEG
para tareas de aprendizaje supervisado.**

Juan Pablo Restrepo Mancilla

Juan Sebastian Rodriguez Solarte

Universidad de Antioquia

Facultad de ingeniería, Bioingeniería

Medellín, Colombia

2022

Introducción

El cerebro es el principal sistema nervioso central. La complejidad e importancia de este órgano hace que la presencia de cualquier tipo de alteración se vea reflejada en el comportamiento de un individuo. Es debido a esta misma importancia que a través de los años se han buscado diferentes técnicas las cuales permitan evaluar el estado del cerebro, y de esta manera detectar y tratar posibles anomalías. Una de las principales ha sido el electroencefalograma, no solo en aplicaciones médicas, sino también ampliamente utilizada en investigaciones relacionadas con la neurología. A pesar de las ventajas que se tienen con uso del electroencefalografías (EEG), como el hecho de ser una técnica no invasiva, posee una desventaja, y es la susceptibilidad que presenta al momento de registrar todo tipo de señales, incluso las que no corresponden a la actividad cerebral, las cuales en muchos casos se podrí­an considera como ruido. Estas señales no deseadas, denominadas artefactos fisiológicos, se combinan con la actividad cerebral, lo que obliga a recurrir a técnicas de separación ciega de fuentes (BSS) como ICA, para realizar el proceso de filtrado. En el caso particular de este trabajo, filtrado espacial.

Planteamiento del problema

Descripción del problema predictivo a resolver. La electroencefalografía (EEG) es una técnica de imagen médica basada principalmente en ubicar electrodos, de manera no invasiva, en la superficie de la cabeza, y de esta forma registrar las señales eléctricas producidas por la actividad cerebral. El trastorno por déficit de atención e hiperactividad (TDAH) es una irregularidad presente en el neurodesarrollo infantil, la cual se puede seguir presentando en la edad adulta. Debido a lo anterior se requieren herramientas que soporten, a temprana edad, un diagnóstico adecuado de este trastorno, para poderlo tratar de manera eficiente. Más aún cuando el TDAH tiene un alto nivel de prevalencia en Colombia, específicamente en Antioquia y el Eje Cafetero (15 % a 17 %) Usando los datos obtenidos de EEG para 9 sujetos, a los que se les pidió pensar entre derecha o izquierda, se entrenará una máquina con validación cruzada para cada uno de los sujetos y así lograr que la máquina prediga si el sujeto pensaba en derecha o izquierda. El software será vendido a una entidad prestadora de salud, en específico al área psiquiátrica. Se busca aumentar la cantidad de pacientes interesados en realizar la prueba, en caso de que los resultados sean satisfactorios pero la cantidad de pacientes que realicen la prueba no aumente, se vería en ese caso, cómo aumentar el alcance y el conocimiento general acerca de la prueba y sus beneficios.

Objetivos

General Analizar las técnicas de separación de artefactos fisiológicos producidos por el movimiento ocular de las señales de EEG utilizando métodos estadísticos basados en la separación ciega de fuentes, como lo son ICA, con el fin de mejorar el desempeño de tareas de clasificación, a partir de series de tiempo.

Específicos Evaluar el desempeño en tareas de clasificación de señales EEG de descriptores estadísticos para variables aleatorias como estrategia de selección de fuentes espaciales. Analizar metodologías supervisadas de selección de fuentes espaciales para mejorar el rendimiento de esquemas de clasificación de señales EEG.

Estudio del Dataset

El dataset fue obtenido de "Competencia BCI IV", consiste en validar el procesamiento de señales y los métodos de clasificación para las interfaces cerebro-computadora. El dataset consta de 9 matrices, cada matriz corresponde a un sujeto diferente y cada una de las matrices consta de 3 dimensiones. Una de las dimensiones es el número de puntos en el tiempo, que para nuestro caso será de 500 y este equivale al número de filas, 22 canales, que equivalen al número de columnas y esto corresponde al número de columnas y 288 trials, que corresponde a la profundidad de la matriz y equivale al número de repeticiones que se hizo el experimento con cada sujeto.

Análisis Dimensional

Se descargó el dataset y se comenzó a analizar sus características, la forma de almacenamiento en memoria de estos es usando diccionarios, cada uno con sus claves y comenzamos con el análisis dimensional de las señales que se estudiaron.

```
#cargamos el dataset
mat_contents = sio.loadmat('NEW_22ch_A09.mat')

#los datos se cargan como un diccionario, se puede evaluar los campos que contiene
print("Los campos cargados son: " + str(mat_contents.keys()))

#queremos extraer solamente la señal eeg
data=sio.loadmat("NEW_22ch_A09.mat")["X"]
#Mostramos las dimensiones de la señal
dims_control = [sio.loadmat("NEW_22ch_A09.mat")["X"].shape ]
```

Figura 1. Cargamos el dataset como diccionarios y claves.

Para una mejor comprensión de los datos decidimos graficar solamente una señal y de esta manera entender cómo funcionan individualmente cada una, y posteriormente una maro cantidad de estas señales en una sola gráfica.

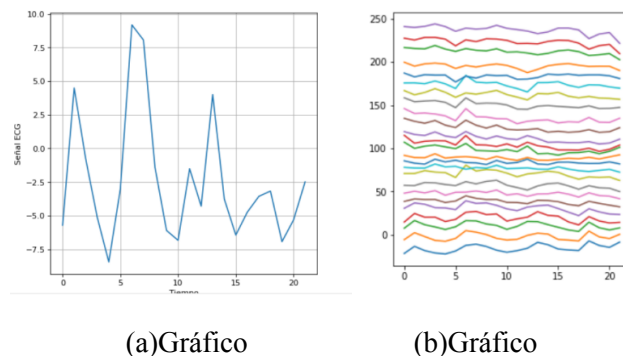


Figura 2. Graficamos las señales individualmente(a) y en conjunto varias señales (b).

Luego de ver individualmente las señales se decidió realizar un programa con interfaz gráfica (Figura 3) que nos permitiera recorrer más fácilmente toda la señal de diferentes personas, por tanto se procedió a realizarlo mediante el paradigma orientado a objetos y utilizando MVC(modelo vista controlador), aparte de ello se sacaron parámetros importantes en la señales como máximos , mínimos, desviación estándar,valor cuadrático medio (RMS)Sample entropy, que se utiliza para evaluar la complejidad de las señales fisiológicas de series temporales y diagnosticar estados patológicos.

Estas características nos van a servir a la hora de crear el modelo de machine learning, por otra parte un reto que tenemos es identificar cuales son las partes de la señal que nos aportan más información.

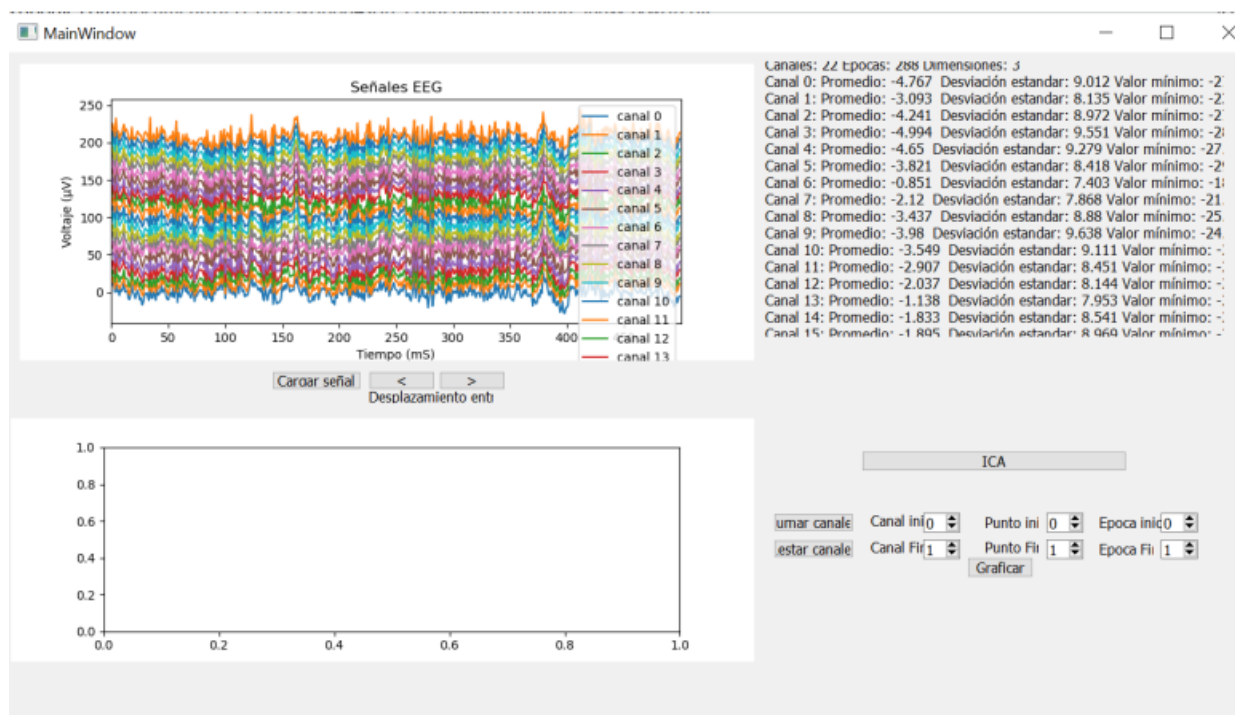


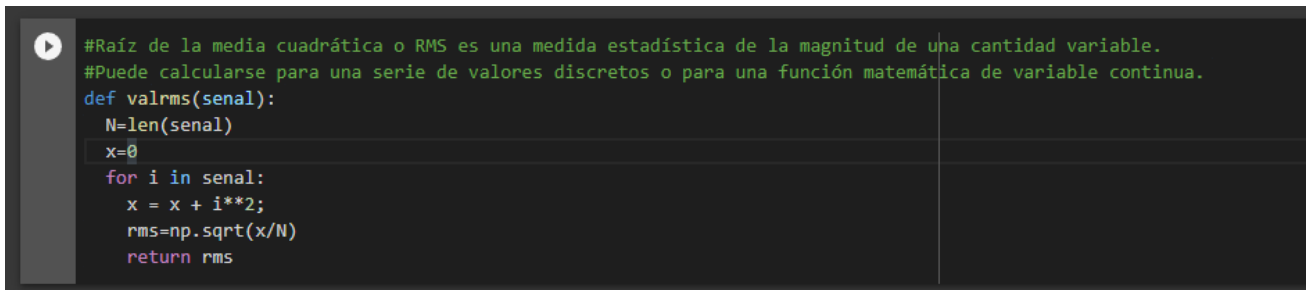
Figura 3. Interfaz gráfica para el análisis, desplazamiento y manipulación estadística de la base de datos.

Con un mejor entendimiento de los datos se logró una mejor comprensión del reto al cual nos enfrentamos, se deben usar ahora métodos estadísticos buscando dentro de espacios multidimensionales los componentes principales, se pretende que si las señales originales son estadísticamente independientes, las señales recuperadas también deben de serlo.

Filtrado y procesamiento estadístico

La idea principal de este documento es mostrar de diferentes formas los resultados obtenidos cuando se aplicaron técnicas de filtrado espacial en distintos electroencefalogramas, con el fin de plantear una comparación entre técnicas supervisadas y no supervisadas de remoción de señales no deseadas, y de esta manera buscar ventajas y desventajas de cada una.

Media cuadrática: El primer método estadístico para el estudio de las señales que se decidió emplear fue la media cuadrática, la cual es una medida de posición central de la estadística descriptiva. La media cuadrática es igual a la raíz cuadrada de la media aritmética de los cuadrados de los datos. Es muy útil para calcular la media de variables que toman valores negativos y positivos. Se suele utilizar cuando el símbolo de la variable no es importante y lo que interesa es el valor absoluto del elemento. Por ejemplo, para calcular la media de errores de medida.



```
#Raíz de la media cuadrática o RMS es una medida estadística de la magnitud de una cantidad variable.
#Puede calcularse para una serie de valores discretos o para una función matemática de variable continua.
def valrms(senal):
    N=len(senal)
    x=0
    for i in senal:
        x = x + i**2;
    rms=np.sqrt(x/N)
    return rms
```

Figura 4. Media cuadrática aplicada a los datos.

ICA Y KURTOSIS

Reduciendo la cantidad de ruido ocular presenten en las bases de datos (EEG), primero utilizando ICA para obtener una buena estimación de las fuentes espaciales, y luego eliminando aquellas con mayor contenido de artefactos oculares. En el procesamiento de señales, el **análisis de componentes independientes** (ICA) es un método computacional para separar una señal **multivariante en subcomponentes aditivos**. Esto se hace suponiendo que como máximo un subcomponente es gaussiano y que los subcomponentes son estadísticamente independientes entre sí. La independencia de las señales es importante para que este método funcione. (Stewart et al., 2014).

Kurtosis . Es una **medida del grado de cuántos valores atípicos hay**. Para una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que los datos muestran más valores atípicos extremos que una distribución normal. Una kurtosis alta es un indicador de la existencia de valores bastante alejados de la media, o en este caso, de posibles señales que no corresponden a la actividad cerebral. (Udmale & Singh, 2019).

Técnicas de Machine learning

Una vez obtenida las señales debidamente filtradas se procedió a aplicar en ellas diferentes técnicas de machine learning. Obteniendo los mejores resultados con las tres siguientes.

1. Linear Discriminant Analysis

análisis discriminante es una técnica predictiva de clasificación ad hoc y se denomina así por que se conocen previamente los grupos o clases antes de realizar la clasificación, que a diferencia de los árboles de decisión (post hoc) donde los grupos de clasificación se derivan de la ejecución de la técnica misma sin conocerse previamente. Es una técnica ideal para construir un modelo predictivo y pronosticar al grupo o clase a la que pertenece una observación a partir de determinadas características que delimitan su perfil. (Feldesman, 2002).

2. Support vector machine

En el aprendizaje automático, las máquinas de vectores de soporte son modelos de aprendizaje supervisado con algoritmos de aprendizaje asociados que analizan datos para clasificación y análisis de regresión. Las SVM son uno de los métodos de predicción más robustos, basados en marcos de aprendizaje estadístico o teoría. Dado un conjunto de ejemplos de entrenamiento, cada uno marcado como perteneciente a una de dos categorías, un algoritmo de entrenamiento SVM construye un modelo que asigna nuevos ejemplos a una categoría u otra, convirtiéndose en un clasificador lineal binario no probabilístico. SVM asigna ejemplos de entrenamiento a puntos en el espacio para maximizar el ancho de la brecha entre las dos categorías. Luego, los nuevos ejemplos se mapean en ese mismo espacio y se predice que pertenecen a una categoría según el lado de la brecha en el que se encuentran. (Noble, 2006).

3. Bayer

Se utiliza una versión simplificada del Teorema de Bayes, conocida como Clasificación Naive Bayes, para reducir el tiempo y los costos de cálculo. En este artículo, lo llevamos a cabo a través de estos conceptos y discutimos las aplicaciones del Teorema de Bayes en el aprendizaje automático.

El teorema de Bayes es un método para determinar las probabilidades condicionales, es decir, la probabilidad de que ocurra un evento dado que ya ocurrió otro evento. Debido a que una probabilidad condicional incluye condiciones adicionales, en otras palabras, más datos, puede contribuir a obtener resultados más precisos.

Por lo tanto, las probabilidades condicionales son imprescindibles para determinar predicciones y probabilidades precisas en el aprendizaje automático. Dado que el campo se está volviendo cada vez más omnipresente en una variedad de dominios, es importante comprender el papel de los algoritmos y métodos como el teorema de Bayes en el aprendizaje automático.

```
#Ciclo del número de componentes a filtrar
for n_ica in range(20):
    #Filtramos todos los trials según MAC
    Xf = np.array([KUR(S[trial], A[trial], means[trial], n_ica) for trial in range(len(S))])
    Xf = np.transpose(Xf, [0, 2, 1])
    #Ciclo de número de componentes en CSP
    for n_components in range(2, Xf.shape[1], 2):
        #Ciclo de validación cruzada
        acc=[]
        acc2=[]
        acc3=[]
        for tr, val in skf.split(labels, labels):
            csp = CSP(n_components=n_components).fit(Xf[tr], labels[tr])
            features = csp.transform(Xdata)

            #linear discrimination
            clf=LDA().fit(features[tr], labels[tr])
            acc.append(clf.score(features[val], labels[val]))
            #g=clf.predict(features[tr])

            #Bayes
            clf = GaussianNB()
            clf=clf.fit(features[tr], labels[tr])
            acc2.append(clf.score(features[val], labels[val]))
            #g3=clf.predict(features[tr])

            # support vector machine
            clf = svm.SVC()
            clf=clf.fit(features[tr], labels[tr])
            acc3.append(clf.score(features[val], labels[val]))
            # g2=clf.predict(features[tr])

        cv_results.append({'n_ica':n_ica, 'n_components':n_components, 'scores':acc, 'mean_score':np.mean(acc), 'std_score':np.std(acc)})
        cv_results2.append({'n_ica':n_ica, 'n_components':n_components, 'scores':acc2, 'mean_score':np.mean(acc2), 'std_score':np.std(acc2)})
        cv_results3.append({'n_ica':n_ica, 'n_components':n_components, 'scores':acc3, 'mean_score':np.mean(acc3), 'std_score':np.std(acc3)})
```

Figura 5. Diferentes técnicas de machine learning.

CURVA DE APRENDIZAJE

Posteriormente se decidió reducir la mayor cantidad ruido posible aplicando una técnica de filtrado de señales que se puede ver en la imagen, está demostró, como se verá más adelante, ser efectiva y obtener mejores resultados en aquellas muestras en las que se usó. Se puede observar el primer sujeto que, sin realizar ningún tipo de filtrado, tuvo un desempeño menor al 90 %, y a pesar de que usando cualquier técnica se llega a superar este desempeño, es más evidente como en este caso las técnicas eliminaron información importante más rápidamente.

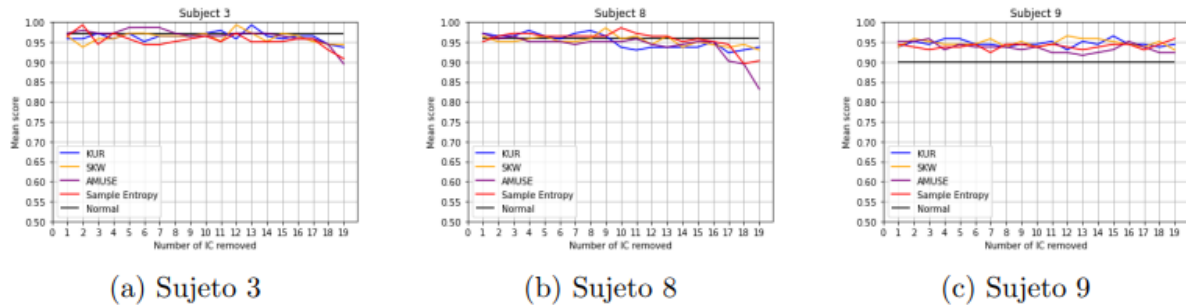


Figura 6. Comparativa de sujetos cuyas ensayos fueron sometidos a filtración. Sujeto C obtuvo los mejores resultados, notan dos cambios significativos respecto a los sujetos anteriores. El primero de ellos es que el acierto medio

Análisis de resultados y conclusiones.

Una vez filtrados los datos usando métodos de ICA y Kurtosis, conociendo bien la estructura, tipo y naturaleza de los datos se aplicaron los tres métodos de machine learning, para los cuales obtuvimos resultados diferentes, estos resultados varían enormemente dependiendo de las técnicas de filtrado ya que el mayor reto de este estudio fue filtrar las señales de la mejor manera, por tanto con estas específicas técnicas de filtrado se obtuvieron los siguientes resultados.

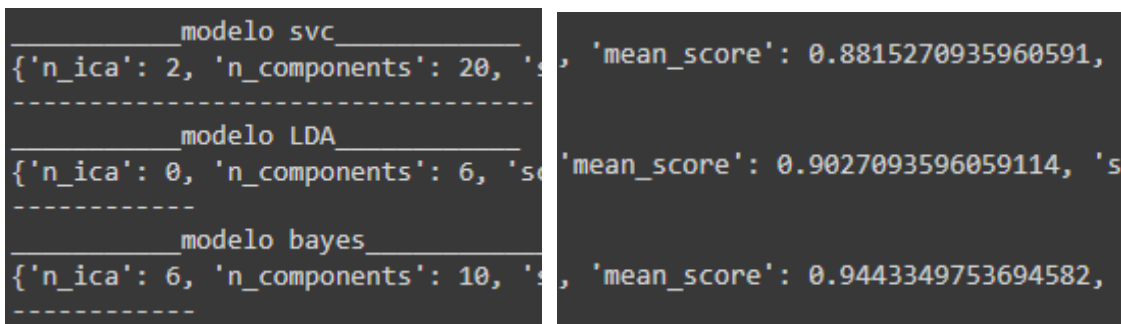


Figura 7. Resultados obtenidos de cada uno de los métodos de machine learning con sus respectivos resultados.

Finalmente podemos observar en la figura 7 que el método con mejores resultados con 94.44% de media es el modelo de bayes, esto se puede deber a la naturaleza de los datos o a ser un método sumamente robusto para conseguir estos resultados.

A partir de estos resultados satisfactorios en las predicciones se podría aspirar a encontrar métodos de machine learning con una mejor efectividad y conseguir unos resultados aún mejores o mejores técnicas de filtración.

Bibliografía

- [1] Stewart, A. X., Nuthmann, A., & Sanguinetti, G. (2014). Single-trial classification of EEG in a visual object task using ICA and machine learning. *Journal of Neuroscience Methods*, 228, 1–14. <https://doi.org/10.1016/j.jneumeth.2014.02.014>
- [2] Udmale, S. S., & Singh, S. K. (2019). Application of Spectral Kurtosis and Improved Extreme Learning Machine for Bearing Fault Classification. *IEEE Transactions on Instrumentation and Measurement*, 68(11), 4222–4233. <https://doi.org/10.1109/tim.2018.2890329>
- [3] Pushpa, G., & Palani, S. (2012). Brain tumor MRI image classification with feature selection and extraction using linear discriminant analysis. ArXiv.org. <https://doi.org/10.48550/arXiv.1208.2128>
- [4] Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, 24(12), 1565–1567. <https://doi.org/10.1038/nbt1206-1565>.
- [5] Por, I., & Karim Paz. (n.d.). MEDIA ARITMÉTICA SIMPLE. <http://www.editorialkamar.com/et/archivo11.pdf>
- [6] D. A. Pineda, F. Lopera, J. D. Palacio, D. Ramirez, and G. C. Henao, “Prevalence estimations of attention-deficit/hyperactivity disorder: Differential diagnoses and comorbidities in a colombian sample,” *International Journal of Neuroscience*, vol. 113, no. 1, pp. 49–71, 2009.

