

Desarrollo de un modelo fundacional estocástico basado en procesos gaussianos para la clasificación de bioseñales EEG en el diagnóstico asistido del TDAH.

Julián David Pastrana Cortés, M.Sc.

Índice

1. Planteamiento del problema y pregunta de investigación	1
2. Justificación	3
3. Estado del arte	3
4. Objetivos	3
5. Metodología	3

Resumen

Put an abstract here!!!

1. Planteamiento del problema y pregunta de investigación

Los trastornos mentales son perturbaciones que afectan la cognición, el comportamiento y las emociones de los individuos, con una prevalencia estimada en aproximadamente 350 millones de personas a nivel mundial [1].

Entre estos trastornos, el Trastorno por Déficit de Atención e Hiperactividad (TDAH) afecta al 5 % de los niños y adolescentes y al 2.5 % de los adultos, lo que incrementa el riesgo de padecer otros trastornos psiquiátricos, dificultades en el ámbito educativo y laboral, así como una mayor propensión a conductas delictivas y adicciones [2].

La detección temprana del TDAH es fundamental para proporcionar un tratamiento oportuno al individuo. No obstante, muchas de las pruebas

diagnósticas actuales son tediosas, requieren un seguimiento prolongado del paciente y están sujetas a interpretaciones subjetivas. Además, en muchas regiones, el acceso a especialistas es limitado (**citar**).

El desarrollo de modelos de aprendizaje automático ha permitido la creación de herramientas que facilitan y respaldan el diagnóstico de diversas enfermedades, incluido el TDAH, lo que contribuye a reducir la carga sobre los especialistas (**citar ejemplos**). La mayoría de estos modelos emplean señales electroencefalográficas (EEG) para clasificar registros como pertenecientes a individuos con o sin la enfermedad, en un enfoque de clasificación binaria. Sin embargo, las señales EEG son altamente no estacionarias y presentan patrones complejos que pueden dificultar su interpretación y manejo mediante modelos tradicionales (**citar**).

Una alternativa prometedora es el uso de modelos basados en aprendizaje profundo, los cuales emplean una gran cantidad de parámetros para extraer características abstractas de las señales EEG que faciliten su clasificación (**citar ejemplos**). No obstante, el entrenamiento de estos modelos requiere una gran cantidad de datos etiquetados dentro de un marco de aprendizaje supervisado. En muchos escenarios, la recolección y el etiquetado de un conjunto de datos de gran tamaño resulta costosa o incluso inviable.

Los modelos fundacionales han surgido como una solución a este problema, ya que pueden ser preentrenados con conjuntos de datos no etiquetados en una fase inicial y luego ajustados para tareas específicas con una cantidad significativamente menor de datos etiquetados (**citar**).

En diversos contextos clínicos, no solo es importante obtener una predicción puntual del diagnóstico, sino también una medida de la incertidumbre asociada a dicho pronóstico. En este sentido, los procesos gaussianos representan una alternativa viable, ya que proporcionan distribuciones de probabilidad sobre las predicciones en lugar de valores deterministas. Los procesos gaussianos han sido ampliamente utilizados para modelar funciones complejas y realizar tareas de clasificación (**citar estudios**). Sin embargo, su integración con modelos fundacionales aún no ha sido ampliamente explorada.

Además, en la fase de preentrenamiento, los conjuntos de datos no etiquetados suelen presentar variabilidad en términos de estándares de adquisición, lo que puede traducirse en valores ausentes, diferencias en la tasa de muestreo o variaciones en el número de canales, entre otros factores. Descartar estos datos podría implicar la pérdida de información valiosa. Por lo tanto, un modelo fundacional diseñado para el análisis de señales EEG debe ser capaz de manejar datos con características ausentes de manera efectiva.

2. Justificación

3. Estado del arte

4. Objetivos

Objetivo General

Desarrollar un modelo de clasificación de bioseñales EEG basado en un modelo fundacional, que integre bases de datos etiquetadas y no etiquetadas mediante técnicas de autoaprendizaje y ajuste fino.

Objetivos específicos

1. Diseñar y desarrollar un modelo fundacional para la clasificación de bioseñales EEG, que aproveche datos no etiquetados en la etapa de autoaprendizaje y datos etiquetados para su ajuste fino.
2. Implementar una herramienta de predicción estocástica basada en procesos Gaussianos, que permita modelar la incertidumbre en la predicción del modelo fundacional.
3. Elaborar una estrategia para el manejo de características faltantes que complemente y extienda el modelo fundacional para tratar bioseñales con información faltante.

5. Metodología

Referencias

- [1] Mohamad Dehghan-Bonari, Mohammad Alipour-Vaezi, Mohammad Mahdi Nasiri, and Amir Aghsami. A diagnostic analytics model for managing post-disaster symptoms of depression and anxiety among students using a novel data-driven optimization approach. *Healthcare Analytics*, 4, 12 2023.
- [2] Stephen V. Faraone, Philip Asherson, Tobias Banaschewski, Joseph Biederman, Jan K. Buitelaar, Josep Antoni Ramos-Quiroga, Luis Augusto Rohde, Edmund J.S. Sonuga-Barke, Rosemary Tannock, and Barbara Franke. Attention-deficit/hyperactivity disorder. *Nature Reviews Disease Primers*, 1, 2015. Cited by: 1053.