

CLASIFICACIÓN DE LA ENFERMEDAD DE ALZHEIMER

Proyecto fundamentos de Deep Learning

CONTEXTO

La enfermedad de Alzheimer es un desorden neurodegenerativo en el cual la actividad cerebral comienza a ser más lenta de lo que es en personas sanas. Las formas genéticas de esta enfermedad se caracterizan por una presentación de los síntomas más temprana, y también en estos casos, se ve un curso más agresivo y rápido en la progresión. Por lo anterior, el estudio de las formas genéticas de esta enfermedad representa una ventaja a la hora del diagnóstico en comparación con las formas esporádicas, convirtiéndola en un modelo único para el estudio de enfermedades neurodegenerativas [1, 2].

Los abordajes para el diagnóstico de la enfermedad de Alzheimer (EA) se han centrado en la descripción clínica ya sea a través de la evaluación neurológica o por medio de pruebas neuropsicológicas, y se ha demostrado como los cambios en el funcionamiento cerebral antecede a los cambios clínicos, por lo que identificar los cambios preclínicos de la enfermedad podría suponer la posibilidad de lograr un diagnóstico temprano. El diagnóstico temprano de la EA supone la facilidad en la implementación de terapias y tratamientos de manera oportuna lo que puede permitir disminuir el avance de atrofia cerebral [1, 2].

Diferentes técnicas han identificado posibles biomarcadores de la EA, entre los cuales se encuentran las neuroimágenes como las imágenes por resonancia magnética o tomografía por emisión de positrones. Una de las técnicas que ha cobrado importancia en el estudio de la actividad cerebral en los últimos años, es el registro encefalográfico (EEG). Actualmente hay estudios en los que han usado las medidas espectrales y la sincronización entre las regiones cerebrales para la búsqueda de biomarcadores en EEG. Los pacientes con EA por lo general muestran una actividad oscilatoria reducida y también se ha observado reducción de la potencia en algunas bandas espectrales [1, 2, 3].

El EEG, por ser una técnica portable y de bajo costo, facilita la masificación de la información y la obtención de grandes volúmenes de datos posibilitando la aplicación de técnicas de “Big Data” que podrían abrir nuevos caminos para el estudio de la EA. Sin embargo, como el uso de EEG para el diagnóstico del alzhéimer aún no se comprobado como método principal, es necesario el uso de información adicional, así como la medida de deposición de amiloides en el líquido cefalorraquídeo o información clínica proveniente de estudios de evaluación neurológica.

Por lo anterior, el objetivo de este proyecto es poder crear un modelo de Deep Learning que permita la clasificación de individuos con riesgo de demencia tipo Alzheimer a partir de bases de datos de EEG e información clínica.

OBJETIVO DEL MODELO

El objetivo del modelo es poder hacer una clasificación de tipos de deterioro cognitivo y controles en un conjunto de datos de EEG. La idea de usar un algoritmo de Deep Learning como lo podría ser una red convolucional que pueda extraer características específicas de cada señal.

DATASET: TIPO DE DATOS, TAMAÑO (NÚMERO DE DATOS Y TAMAÑO EN DISCO), DISTRIBUCIÓN DE LAS CLASES

El conjunto de datos que se planea usar en este proyecto es una base de datos provenientes de registros de EEG realizados a personas portadoras de un gen que promueve el desarrollo de Alzheimer y pacientes control (sanos no portadores). El dataset está compuesto por información extraída de la señal EEG como la potencia o la entropía, y una columna asociada al grupo, que se refiere a si el sujeto es portador o control, esta columna será la que presuntamente se usará como salida del modelo.

MÉTRICAS DE DESEMPEÑO (DE MACHINE LEARNING Y NEGOCIO)

Las métricas de machine learning que se pretenden utilizar en este proyecto son la precisión, sin embargo, esta se debe acompañar al menos por otra métrica más. En la literatura se ha observado que, para problemas de clasificación con varias clases, similares al que vamos a tratar en este proyecto, es útil usar una matriz de confusión y una curva ROC, ya que los datos pueden ser desbalanceados y además se quiere tener conocimiento de la sensibilidad del modelo.

Por otro lado, la idea del modelo planteado en este proyecto tiene que ver con la efectividad de los tratamientos gracias a un diagnóstico temprano del Alzheimer o demencia leve. Para poder determinar la efectividad del modelo obtenido en términos de este “negocio”, se debe evaluar medicamente si al abordar en etapas tempranas un paciente que fue clasificado como una persona con potencial desarrollo del Alzheimer, el deterioro cognitivo si disminuye, sin embargo, obtener esta medida es complicado debido a que se requiere de mucho tiempo para obtener esa conclusión y, además, se necesita un constante acompañamiento clínico a lo largo de los años. Por lo tanto, por efectos de practicidad, se puede plantear la evaluación del modelo usando datos adquiridos de personas portadoras de gen PSEN1, las cuales, según estudios realizados por el grupo de neurociencias de Antioquia (GNA), está comprobado que adquirirán Alzheimer.

Otra métrica de desempeño que podría usarse sería estudiar la disminución de gastos y tiempo que habría si se clasifica adecuadamente una persona con Alzheimer.

RESULTADOS PREVIOS

En estudios recientes, se ha encontrado que el EEG puede servir como potencial método para la detección temprana de Alzheimer. Algunos enfoques están basados en técnicas de inteligencia artificial para la clasificación de esta enfermedad con grandes volúmenes de datos, e incluso con datasets que no son tan extensos.

En la universidad de Palermo [4] se hizo una revisión de la literatura con el objetivo de demostrar que los modelos automatizados pueden diferenciar y clasificar el Deterioro Cognitivo Leve (DCL) y la enfermedad de EA. Dentro de los artículos que se revisaron, se encontró los algoritmos más utilizados para la clasificación son: SVM y Redes Neuronales, pero el algoritmo que tuvo mejor precisión fue NB, con precisiones de hasta un 98%.

Por otro lado, se ha demostrado que, en este enfoque, el diseño de modelos de Deep Learning multi tarea para reducir el problema de sobre ajuste, muestra una mejor capacidad de extracción de representaciones de alto nivel lo que representa un avance en el estado del arte [5].

Así mismo en otros estudios, se han experimentado nuevos enfoques y diseños de modelos de Deep Learning para superar los obstáculos relacionados con la disponibilidad de datos y el problema de sobreajuste. Estos métodos han demostrado mejoras significativas en la exactitud y precisión de los modelos con respecto a otros métodos de Machine Learning [4, 5, 6].

Por lo anterior, se espera obtener un modelo de Deep Learning usando los conceptos vistos en el curso que permita la clasificación de un conjunto de datos entre portadores y no portadores de un gen específico que provoca Alzheimer o demencia leve. Se espera que las métricas del modelo sean satisfactorias y comparables con las de la literatura con una precisión del al menos un 65%.

REFERENCIAS

- [1] Rossini, P. M., di Iorio, R., Vecchio, F., Anfossi, M., Babiloni, C., Bozzali, M., Bruni, A. C., Cappa, S. F., Escudero, J., Fraga, F. J., Giannakopoulos, P., Guntekin, B., Logroscino, G., Marra, C., Miraglia, F., Panza, F., Tecchio, F., Pascual-Leone, A., & Dubois, B. (2020). Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts. *Clinical Neurophysiology*, 131(6), 1287–1310. <https://doi.org/10.1016/J.CLINPH.2020.03.003>
- [2] Gaubert, S., Raimondo, F., Houot, M., Corsi, M. C., Naccache, L., Sitt, J. D., Hermann, B., Oudiette, D., Gagliardi, G., Habert, M. O., Dubois, B., de Vico Fallani, F., Bakardjian, H., & Epelbaum, S. (2019). EEG evidence of compensatory mechanisms in preclinical Alzheimer's disease. *Brain*, 142(7), 2096–2112. <https://doi.org/10.1093/BRAIN/AWZ150>
- [3] Rodríguez, G., & Díaz, V. (2020). *Aprendizaje automático para detección de problemas cognitivos: una revisión de la literatura (Machine Learning for detection of cognitive problems: a review of the literature)*. 9–22.
- [4] Kim, M., Youn, Y. C., & Paik, J. (2023). Deep learning-based EEG analysis to classify normal, mild cognitive impairment, and dementia: algorithms and dataset. *NeuroImage*, 120054. <https://doi.org/10.1016/J.NEUROIMAGE.2023.120054>
- [5] Bi, X., & Wang, H. (2019). Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning. *Neural Networks*, 114, 119–135. <https://doi.org/10.1016/J.NEUNET.2019.02.005>
- [6] Xia, W., Zhang, R., Zhang, X., & Usman, M. (2023). A novel method for diagnosing Alzheimer's disease using deep pyramid CNN based on EEG signals. *Heliyon*, 9(4), e14858. <https://doi.org/10.1016/J.HELİYON.2023.E14858>