УДК 519.85

Гайдар В.О.¹, аспірант

**Аналіз методів машинного навчання для
раннього детектування та прогнозування
епілептичних нападів.**

¹ Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4г,
e-mail: gaidar.viktoria@gmail.com

Gaidar V.O.¹, Postgraduate student

**Machine learning for epilepsy detection and
forecast review: new challenges and
perspectives.**

¹ Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4g,
e-mail: gaidar.viktoria@gmail.com

Проведено порівняльний аналіз методів машинного навчання для вирішення задачі раннього прогнозування та діагностики епілептичних нападів за допомогою електроенцефалографічних сигналів. Побудована математична модель досліджуваних алгоритмів та наведено кількісні дані для оцінки ефективності методів.

Ключові слова: машинне навчання, епілепсія, електроенцефалографія, передіктальні стани.

The comparative analysis of machine learning methods has performed to solve the problem of early detection and prediction of epileptic seizures using electroencephalographic signals. Recent studies has shown that it is possible to predict seizures in prior of its physical appearance. Our goal is to present and analyse different approaches of seizure prediction techniques, particularly in machine learning and deep learning. Seizure prediction has made important advances over the last decade, nevertheless it is still a problem to provide steady algorithm of seizure early detection. Also, within individual patients exhibit distinctive dynamics, is it crucial to find algorithms providing greater clinical utility. This article focuses of the problem of features development from electroencephalography signals in order to provide the accurate pattern recognition techniques for detection and classification of epilepsy seizures in advance. The mathematical model of the algorithms is constructed and quantitative data presented for estimating the methods efficiency.

Key Words: machine learning, epilepsy, electroencephalogram, preictal seizures.

Статтю представив д.ф.-м.н., проф. Анісімов І.О.

Вступ

Епілепсія – це стан, що характеризується функціональним порушенням роботи головного мозку. Непередбачуваний характер нападів, що є характерним проявом захворювання, становить серйозну загрозу для життя та здоров'я людини. Для більшості випадків причина хвороби невідома, тому задача ранньої діагностики та прогнозування нападів наразі активно вивчається. Ефективним методом графічної реєстрації біоелектричної активності головного мозку, що використовують при дослідженні епілепсії є електроенцефалографія (ЕЕГ). У клінічних дослідженнях використовують багатоканальні (від ста двадцяти шести електродів) відведення та середній час одного

дослідження триває від декількох годин до трьох діб. Відповідно, виникає необхідність в аналізі великих масивів даних у режимі реального часу. З розвитком методів машинного навчання, необхідність оптимізації виділення основних статистичних характеристик ЕЕГ сигналу відіграє вирішальну роль при створенні алгоритмів прогнозування на їх основі. Метою даної роботи є визначення основних підходів та методів, що використовуються для вирішення проблеми класифікації та виділення комплексів пік-хвилі та спайкової активності у передіктальний період. Проаналізувати основні статистичні характеристики ЕЕГ сигналів, що слугують вхідними параметрами до

запропонованих методів машинного навчання та провести оцінку їх ефективності.

Передумови раннього детектування епілептичних нападів

Електроенцефалографія є однією з простих та мобільних методів для виявлення аномалій, які спостерігаються при епілепсії. Цей метод діагностики безпечний і багатофункціональний для вирішення задач розпізнавання, локалізації і виявлення синдрому епілепсії [1]

Рис. 1 Стандартна схема підключення електроенцефалографу «10-20 %»

У даній роботі порівнюються підходи машинного навчання до побудови специфічних до пацієнта класифікаторів, які дозволяють передбачити епілептичний напад за допомогою таких методів як згорткові нейронні мережі (ЗНН), байєсівський класифікатор (БК), метод k -найближчих сусідів (k -НС) і метод опорних векторів (МОВ). Розглядаються статистичні особливості, які найкраще вирішують проблему виділення патернів і дають достовірні результати. Поставлено деякі питання, можливості вирішення проблеми прогнозування епілепсії в термінах простої класифікаційної задачі.

На сьогоднішній день було доведено існування перед епілептичних станів, що з'являються напередодні фізичного нападу. Ці стани залежать не тільки від типу епілепсії, а також і від особливостей окремого пацієнта. На даний момент не вирішеною залишається проблема виділення тих компонентів, які могли б описати зміни у передіктальний період як для окремого так і для групи пацієнтів.

Виділення статистичних характеристик

Прогнозування нападів є основною сферою досліджень, які базуються на аналізі особливостей ЕЕГ записів. Проте в даний час не існує надійних алгоритмів прогнозування нападів для клінічних застосувань. Незважаючи на відсутність розуміння механізму іктогенезу епілепсії, дослідження показали існування

патернів, що виникають напередодні судом. Ці клінічні спостереження дають обґрунтування до пошуку специфічних ознак на електроенцефалограмах.

Ранні дослідження [2] намагалися знайти рішення, використовуючи перші кілька періодів ЕЕГ, що передували фізичному нападу. Оскільки, така процедура передбачає, що просторово-часова динаміка не змінюється від захоплення до захоплення, такі алгоритми виявилися ненадійними. Подальші дослідження розробили новий алгоритм, який ітераційно змінює параметри після кожного захоплення.

Одне з перших комплексних досліджень [3] показує, як визначити патерни ЕЕГ для виявлення епілептичних нападів. У цьому дослідженні було побудовано 30 характеристик, включаючи біваріантні і багатоваріантні особливості, а також лінійні та нелінійні параметри. Тут розглянемо лише основні характеристики, що застосовуються при передбаченні. Протягом усього обговорення дотримуються таких позначень: F_k^i – статистична характеристика i -го каналу для пацієнту k .

ЕЕГ дані були записані як стохастична змінна $X = (x_1, x_2, \dots, x_N)$, середнє значення як x_c та дисперсію як $Var(x_1, x_2, \dots, x_N)$ відповідно. Середньоквадратичне відхилення визначено як:

$$\sigma = \frac{1}{N-1} \sum_{i=1}^N [x_i - x]^2 \quad (1)$$

Асиметрія – це кількісна міра порушення симетрії нормального розподілу. Якщо асиметрія позитивна, дані позитивно відхилені або відхилені вправо. Це означає, що правий хвіст розподілу довший за лівий. Асиметрія обчислюється як

$$skewness = \frac{1}{N} \sum_{i=1}^N \left[\frac{x_i - x}{\sigma} \right]^3 \quad (2)$$

Висоту і гостроту центрального піку відносно решти даних вимірюють за допомогою коефіцієнту ексцесу:

$$kurtosis = \frac{1}{N} \sum_{i=1}^N \left[\frac{x_i - x}{\sigma} \right]^4 - 3, \quad (3)$$

Це означає, що форма графіку розподілу ймовірностей більш гостра за нормальний розподіл, якщо вона має позитивне значення і пологіша, якщо вона має від'ємне значення. Використовують енергію сигналу за заданий період [4]:

$$Energy(t, w) = \frac{1}{w} \sum_{i=t-w}^{t-1} s(i)^2, \quad (4)$$

де $s(i)$ – це амплітуда сигналу у момент часу i , значення сигналу відбираються послідовно, починаючи з початкового моменту часу 1 і закінчуючи в момент часу w . Якщо енергія сигналу усереднена по різному часовому вікну, вона розділена на дві характеристики: короточасну енергію (STE) і довгострокову енергію (LTE). Накопичена енергія (AE) є іншим потужним засобом знаходження аномальної поведінки мозку і визначається як сума послідовних значень енергії сигналу у рухомому вікні [5]. AE виражається як:

$$AE(q) = \sum_{i=1}^{t-q} Energy(t-w). \quad (5)$$

Методи дослідження

Для класифікації записів ЕЕГ в один з чотирьох класів – іктальний, перед-іктальний, міжіктальний, пост-іктальний, було досліджено різні методи. Іктальний період відповідає епілептичній активності. Передіктальний – це період, що настає за декілька хвилин до нападу. Постіктальна активність відповідає за період відразу після судом. Міжіктальним називають період між постіктальними і передіктальними патернами. Існує два загальних підходи в дослідженнях: статистичні та алгоритмічні. Статистичний підхід виділяє різні характеристики сигналу ретроспективно, в той час як алгоритмічний підхід позначає кожен момент часу в наборі даних, тому може бути використаний для систем реального часу. Серед методів машинного навчання для вирішення даної задачі використовують наступні алгоритми: згорткові нейронні мережі (ЗНН), байєсівський класифікатор (БК), метод k -найближчих сусідів (k -НС) і метод опорних векторів (МОВ). Розглядаються лише два з вищезгаданих методів – згорткові нейронні мережі та метод опорних векторів. Їх використовують у мобільних електроенцефалографах.

Для виявлення та прогнозування епілепсії використовуються як методи з учителем, так і без учителя. У випадку навчання з учителем найбільш поширеними техніками, є штучні нейронні мережі (ANN). Вони використовуються для виконання певних конкретних завдань, таких як кластеризація, класифікація, розпізнавання образів [6]. Нейронна мережа складається з декількох шарів, метою яких є вилучення різних

рівнів представлення даних, які є актуальними для конкретної проблеми. Фундаментальним елементом обробки нейронної мережі є нейрон. Існує багато варіацій основного типу нейрона, що ускладнює спроби людини електрично відтворювати процес мислення. Проте всі природні нейрони мають чотири основні компоненти – дендрит, сома, аксон і синапс. Дендрити діють як вхідні канали. Соматим обробляє ці вхідні сигнали протягом деякого часу. Далі сома перетворює отриману величину на вихідний сигнал, який посилюється іншим нейронам через аксон і синапси. Нейрон отримує вхідний сигнал різних джерел, об'єднує їх певним чином, виконує нелінійну операцію, а потім виводить кінцевий результат. На рис.2 показано взаємозв'язок цих чотирьох частин [7].

Рис. 2 Схематичне зображення штучного нейрона

Метод опорних векторів – це метод регресійного аналізу даних з керованим навчанням. Основною задачею є створення алгоритму класифікації даних за допомогою навчальної вибірки, яка вже має поділ на класи. Він використовується для класифікації нових значень. Алгоритм має два основні етапи: відображення даних у простір ознак, за допомогою ядра методу та побудова розділової гіперплощини.

Добірка даних за допомогою нелінійного відображення «переноситься» у простір з більшою розмірністю, що називається простором ознак. Ядро це функція $K(x, y)$, що відображає вибірку даних у множину дійсних чисел R . Існує багато видів ядер, які застосовуються на практиці: поліноміальне [8], гауссове, експотенціальне. Для побудови розділової гіперплощини необхідно враховувати лінійну роздільність набору даних. Для обох випадків є стандартний алгоритм визначення розділової гіперплощини. $Y \in [-1; 1]$ вказує на клас до якого належить точка x_i , що є r -вимірним вектором у просторі ознак. Необхідно знайти таку гіперплощину, яка чітко розмежовує x_i для яких $y_i = 1$ та x_i для яких $y_i = -1$.

Ця гіперплощина максимально рівновіддалена від обох груп даних. Необхідно обрати дві паралельні гіперплощини, які розділяють два класи даних таким чином, щоб відстань між ними була максимальною. Те саме справедливо і для задачі з декількома класами.

Висновки

Аналіз існуючих алгоритмів ранньої діагностики та прогнозування епілептичних нападів показав нові перспективи у створенні експертних систем. Методи головних компонент та опорних векторів є ефективними для визначення статистичних властивостей сигналу, що відповідають передіктальним станам. До таких властивостей належать середньоквадратичне відхилення, коефіцієнт асиметрії та параметри вейвлет-перетворення п'ятого порядку. Побудова

штучних нейронних мереж для моделювання динаміки інтракраніальних ЕЕГ сигналів показала можливість виділення патернів епілептичних нападів напередодні їх фізичного прояву. Однак створення мобільних систем передбачає аналіз даних у режимі реального часу. Тому, побудова таких систем потребує вивчення не лише специфічних статистичних характеристик, але і їхньої фізичної природи. Базуючись на проведених дослідженнях, необхідно враховувати фізичні особливості та залучати додаткові фізіологічні параметри, які можуть впливати на розвиток та прояв епілептичних станів та нападів.

Список використаних джерел

1. Alomari M. H. Wavelet-based feature extraction/ Awada, E. A., Samaha, A. - *Electroencephalography and clinical neurophysiology*. – 1975. – P. 435-436
2. Moghim N. Predicting epileptic seizures in advance/ Negin W., PloS one. – 2014.
3. Aschenbrenner-Scheibe R. How well can epileptic seizures be predicted?/ Maiwald, T., Winterhalder, M – *Brain*. – 2003. – 3. 2616-2626
4. Harrison M. A. Accumulated energy revisited/ Mark G. Frei, Ivan Osorio - *Clinical Neurophysiology*. – 2005. – P. 527-531
5. Iasemidis L. D. Long-term prospective on-line real-time seizure prediction/ Shiau, D. S., Pardalos, P. M. – *Clinical Neurophysiology*. – 2015.
6. Viglione S.S. Epileptic seizure prediction/ Walsh G.O. – *Electroencephalography and clinical neurophysiology*. – 1975
7. Rogowski Z. On the prediction of epileptic seizures/ Gath I., Bental E. – *Biological cybernetics*. – 1981 – P. 9-15
8. Frestone D.R. A forward-looking review of seizure prediction/ Karoly P.J., Cook M.J – *Current opinion in neurology*. – 2017 – P. 167-173

References

1. ALOMARI, M.H., AWADA E.A., SAMAHA A. (1975) Wavelet-based feature extraction. *Electroencephalography and clinical neurophysiology*. pp. 435-436
2. MOGHIM N., NEGIN W. (2014) Predicting epileptic seizures in advance. *PloS one*.
3. ASCHENBRENNER-SCHEIBE R., MAIWALD T., WINTERHALDER M. (2003) How well can epileptic seizures be predicted. *Brain*. pp. 2616-2626
4. HARRISON M. A., MARK G. FREI, OSORIO I. (2005) Accumulated energy revisited. *Clinical Neurophysiology*. pp. 527-531
5. IASEMIDIS L.D., SHIAU D.S, ARDALOS P.M. (2015) Long-term prospective on-line real-time seizure prediction. *Clinical Neurophysiology*.. pp. 527-531
6. VIGLIONE S.S., WALSH G.O. (1975) Epileptic seizure prediction. *Electroencephalography and clinical neurophysiology*.
7. ROGOWSKI Z., GATH I., BENTAL E. (1981) On the prediction of epileptic seizures. *Biological cybernetics*. pp. 9-15
8. FRESTON D.R., KAROLY P.J. (2017) A forward-looking review of seizure prediction. *Current opinion in neurology*. pp. 167-173

Надійшла до редколегії 20.11.2018