

Применение методов интерпретируемого машинного обучения для поиска транскриптомных биомаркеров болезни Альцгеймера

Асаинов Д.Т.^{1*}, Кириллов Б.А.^{2,3}

¹ Кафедра биохимии, биотехнологии и фармакологии, Институт фундаментальной медицины и биологии, Казанский (Приволжский) федеральный университет, Казань, Россия

² Центр технологий материалов, Сколковский институт науки и технологий, Москва, Россия

³ Центр высокоточного редактирования и генетических технологий для биомедицины, Институт биологии гена Российской академии наук, Москва, Россия

* dasainov@inbox.ru

Ключевые слова: болезнь Альцгеймера; машинное обучение; биомаркеры; транскриптом

Мотивация и цель: Диагностика болезни Альцгеймера (БА) представляет собой продолжительный процесс [1], включающий в себя изучение анамнеза, тесты, направленные на оценку психического статуса пациента, инвазивные анализы, необходимые для оценки уровня бета-амилоида в спинномозговой жидкости, и методы нейровизуализации. При этом ранняя диагностика с применением классических методов затруднена из-за неоднородности симптомов и их неоднозначного распределения по времени появления, что наряду с необходимостью дифференцирования БА от других видов деменции и невозможностью однозначно подтвердить диагноз до гистологического исследования мозга говорит о серьезном потенциале альтернативных методов диагностики БА. К ним относятся и модели машинного обучения, которые открывают возможности как для высокоточной ранней диагностики, так и для изучения аспектов патогенеза. Однако существует характерная для многих моделей проблема «черного ящика» – иначе говоря, отсутствия информации о том, как модели машинного обучения принимают решения, что ограничивает их применение в клинической практике. Для решения этой проблемы были созданы методы Explainable Machine Learning (xML, IML [2]) – объяснимого или интерпретируемого машинного обучения. Методы xML позволяют получить информацию о механизмах принятия решений и вкладах в прогнозы отдельных входных признаков, что должно повысить уровень доверия к новым для медицинского сообщества способам диагностики и облегчить получение информации о потенциальных драйверах БА.

Методы и алгоритмы: Для обучения моделей мы использовали фреймворк Ludwig [3], ориентированный на методы глубокого обучения. Используемый датасет представляет собой полученные с чипа Illumina Human HT-12 v4 данные об экспрессии РНК в средней височной извилине у пациентов с диагностированной БА и контрольной группы [4]. Для генерации объяснений предсказаний мы использовали алгоритм Accumulated Local Effects (ALE) в составе библиотеки Alibi [5].

Результаты: Мы обучили ряд моделей, прогнозирующих наличие БА, лучшие значения метрик которых составляют: сбалансированная точность – 0.924,

специфичность – 0.886, коэффициент корреляции Мэтьюза – 0.85. Объяснения, полученные с помощью ALE, позволили выделить ряд транскриптов, являющихся особо яркими входными признаками для обученных моделей. Соответствующие гены, представляющие интерес, вовлечены в процессы свертывания крови, возникновения воспалительных реакций, биосинтеза дофамина и поддержания функции макрофагов. В ходе анализа литературы были найдены подтверждения связи этих генов с болезнью Альцгеймера.

Выводы: Результаты нашего исследования демонстрируют возможность извлечения из высокоточных моделей машинного обучения информации о потенциально задействованных в патогенезе БА генах, что потенциально может повысить уровень доверия к подобным моделям в клинической практике и оптимизировать процедуры подтверждения диагноза и ранней диагностики.

Application of Explainable Machine Learning methods to find transcriptomic biomarkers of Alzheimer's disease

Asainov D.^{1*}, Kirillov B.^{2, 3}

¹ Department of Biochemistry, Biotechnology and Pharmacology, Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kazan, Russia

² Materials Technology Center, Skolkovo Institute of Science and Technology, Moscow, Russia

³ Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, Moscow, Russia

* dasainov@inbox.ru

Key words: Alzheimer's disease; machine learning; biomarkers; transcriptome

Motivation and Aim: Diagnosis of Alzheimer's disease (AD) is a lengthy process [1], that includes gathering of anamnesis, tests for the patient's mental status, invasive tests necessary to assess the level of beta-amyloid in the cerebrospinal fluid, and neuroimaging methods. However, early diagnosis using classical methods is complicated by the heterogeneity of symptoms in a cohort of patients with Alzheimer's disease and their ambiguous distribution by time of onset, which, along with the need to differentiate AD from other types of dementia and the inability to confirm the diagnosis before histological examination of the brain, indicates the serious potential of alternative methods of diagnosing AD. Such methods include machine learning models, which offer opportunities both for high-precision early diagnosis and for studying the aspects of pathogenesis. However, the majority of models possess a "black box" problem – the lack of information about how machine learning models make decisions. Such lack of information limits the applicability of machine learning models in clinical practice. Explainable Machine Learning (xML, IML [2]) methods – explainable or interpretable machine learning – have been created to address this problem. xML methods provide information about decision-making mechanisms and contributions of individual input features to the predictions, which should increase the level of confidence in new diagnostic methods for the medical community and facilitate the acquisition of information about potential drivers of AD.

Methods and Algorithms: To train the models, we use the Ludwig framework [3] focused on deep learning methods. The dataset we use is RNA expression data for the middle temporal gyrus of patients with diagnosed AD and controls [4] obtained from the

Illumina Human HT-12 v4 chip. We use the Accumulated Local Effects (ALE) algorithm as part of the Alibi library [5] to generate explanations for our predictions.

Results: We have trained a number of models predicting the presence of AD, the best metrics of which are: Balanced Accuracy – 0.924, Precision – 0.886, Matthews Correlation Coefficient – 0.85. The explanations obtained with ALE allowed us to identify several transcripts that are particularly important input features for the trained models. Relevant genes of interest, found by our ALE test, are involved in blood coagulation, the occurrence of inflammatory responses, dopamine biosynthesis and maintenance of macrophage function. We found literature that supports the association of these genes with Alzheimer's disease.

Conclusion: The results of our study demonstrate the feasibility of extracting information about genes potentially involved in AD pathogenesis from machine learning models, potentially increasing the level of confidence in such models in clinical practice and optimizing diagnosis confirmation and early diagnosis procedures.

Список литературы/References

1. McKhann G.M., Knopman D.S., Chertkow H. et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7(3):263-269
2. Belle V., Papantonis I. Principles and practice of explainable machine learning. *Front Big Data.* 2021;4:688969. doi 10.3389/fdata.2021.688969
3. Molino P., Dudin Y., Miryala S.S. Ludwig: a type-based declarative deep learning toolbox. *arXiv.* 2019. doi 10.48550/arXiv.1909.07930
4. Piras I.S., Krate J., Delvaux E., Nolz J. et al. Transcriptome changes in the Alzheimer's disease middle temporal gyrus: importance of RNA metabolism and mitochondria-associated membrane genes. *J Alzheimers Dis.* 2019;70(3):691-713
5. Klaise J., Looveren A.V., Vacanti G., Coca A. Alibi explain: algorithms for explaining machine learning models. *J Mach Learn Res.* 2021;22(1):8194-8200