

В. Н. Захаров, С. А. Ступников, И. А. Шанин (Москва, ФИЦ ИУ РАН). Математические методы анализа данных электроэнцефалографии с помощью многозадачного обучения.

УДК 004.77+621.37

Резюме: Предложен подход к анализу данных электроэнцефалографии (ЭЭГ) на основе многозадачного обучения искусственных глубоких нейронных сетей. Данный подход позволяет эффективно анализировать неоднородные наборы данных, используя единую нейросетевую архитектуру для решения набора взаимосвязанных задач.

Ключевые слова: нейроинформатика, электроэнцефалография, многозадачное обучение

Анализ сигналов, полученных с помощью электроэнцефалографии позволяет решать множество задач как в области медицины [1][2], так и в области разработки интерфейсов мозг-компьютер [3]. Медицинские применения ЭЭГ включают в себя выявление патологий в работе головного мозга, связанных с эпилепсией, болезнью Альцгеймера, выявление нарушений сна (соннография), а также выявление патологий, вызванных черепно-мозговыми травмами и нарушением кровообращения. В последние годы, благодаря удешевлению технологии и появлению коммерческих устройств, регистрирующих электрические потенциалы, вызванные деятельностью головного мозга, в открытом доступе появилось множество доступных наборов данных [4]. Таким образом, приобретает все большую актуальность задача разработки методов и программных средств анализа большого объема неоднородных наборов данных ЭЭГ [5].

К рассматриваемым авторами задачам относятся распознавание моторно-визуальных активностей, распознавание эмоционального состояния субъекта, построение характеристик, позволяющих распознать пол, возраст и другие биометрические параметры субъекта. Так как перечисленные задачи являются взаимосвязанными, авторами предлагается подход, учитывающий схожесть этих задач при совместном многозадачном обучении. Использование данного подхода позволяет эффективно учитывать опыт, накопленный при решении аналогичных задач для создания единого, более качественного решения.

В зависимости от протокола эксперимента, анализу подлежит либо продолжительный многоканальный электрический сигнал ЭЭГ, либо серия более коротких сигналов, являющихся регистрацией реакции субъекта на серию визуальных или аудиальных стимулов (анализ вызванных потенциалов). Анализ данных ЭЭГ состоит из следующих шагов:

- 1. Предварительная обработка сигнала.** В силу технических особенностей эксперимента, данные ЭЭГ подвержены зашумлению. Для предварительной очистки используются фильтры с конечной импульсной характеристикой, в том числе полосно-заграждающие фильтры.
- 2. Удаление артефактов.** При регистрации электрической активности головного мозга неизбежно регистрируются активности, с мозговой деятельностью

напрямую не связанные. В их число входят движения глаз, мышечная активность шеи и лица, сердечная активность и другие. Для разделения смеси статистически независимых сигналов (и последующего устранения нежелательных составляющих) широко используется анализ независимых компонент (ICA).

3. **Извлечение информативных признаков.** Наиболее информативным представлением сигналов ЭЭГ является частотно-временное представление, получаемое с помощью быстрого преобразования Фурье. Получаемый спектр сигнала принято делить на следующие диапазоны, именуемые в нейрофизиологии ритмами: альфа-ритм (8...13 Гц), бета-ритм (13...30 Гц), гамма-ритм (30...100 Гц) и другие. В ряде работ были показаны значимые зависимости между деятельностью головного мозга и относительной интенсивностью различных ритмов головного мозга.
4. **Применение классификационной или регрессионной модели для построения предсказаний.** Извлекаемые признаки используются как набор независимых переменных, на основе которых обучаются модели машинного обучения, в частности градиентный бустинг и метод опорных векторов (SVM).

Авторами данной работы предложен подход, позволяющий совместить основные шаги анализа данных ЭЭГ в единой нейросетевой архитектуре, состоящей из последовательного применения автокодировщика, сверточного модуля и полносвязной нейронной сети. На вход данная модель принимает признаки, полученные с помощью применения быстрого преобразования Фурье, на основе которых предсказываются значения целевых переменных. Автокодирующая нейронная сеть используется как аналог метода независимых компонент, переводя сигнал в латентное представление меньшей размерности, подходящее для эффективного удаления артефактной информации. Сверточная сеть используется для построения векторного представления частотно-временной информации, используемого далее для предсказания с помощью полносвязной нейронной сети. На этапе анализа векторных представлений применяется несколько независимых полносвязных сетей, каждая из которых используется для предсказания одной зависимой переменной, таким образом данная архитектура являет собой пример многозадачного обучения. Параметры предложенной нейросети получаются в результате совместного обучения, что позволяет добиться дополнительного улучшения качества предсказания.

Работа выполнена при финансовой поддержке РФФИ, проекты 18-07-01434, 18-29-22096.

СПИСОК ЛИТЕРАТУРЫ

1. *Rim B., Sung N.-J., Min S., Hong M.* Deep Learning in Physiological Signal Data: A Survey. *Sensors*, 2020, 20(4):969.
2. *Li G., Lee C.H., Jung J.J., Youn Y.C., Camacho D.* Deep learning for EEG data analytics: A survey. *Concurrency Computaiont Practice and Experience*, 2019, e5199. <https://doi.org/10.1002/cpe.5199>
3. *Zhang X., Yao L., Wang X., Monaghan J., McAlpine D.* A Survey on Deep Learning based Brain Computer Interface: Recent Advances and New Frontiers. *CoRR*, abs/1905.04149, 2019. <http://arxiv.org/abs/1905.04149>
4. *Bryukhov D. O., Stupnikov S. A., Kovalev D. Y., Shanin I. A.* Neurophysiology as a subject domain for data intensive problem solving. — *Informatics and Applications*, 2020, 14(1), p. 40–47.
5. *Kovalev D. Y., Shanin I. A., Tirikov E. M.* Multidisciplinary neuroinformatics problems for execution in distributed computing infrastructures. — *Systems and Means of Informatics*, 2020, 30(2), p. 43–55.

УДК 004.77+621.37

Zakharov V. N., Stupnikov S. A., Shanin I. A. (Moscow, FRCCSC RAS).
Mathematical methods for electroencephalography data analysis using multi-task learning

Abstract: An approach to the analysis of electroencephalography data based on deep multi-task learning is proposed. This approach allows efficient analysis of heterogeneous datasets using a single neural network architecture to solve a set of interrelated tasks.

Keywords: neuroinformatics, electroencephalography, multi-task learning