

Классификация данных ЭЭГ с помощью WEKA

М.Ш. Муртазина¹, Т.В. Авдеенко¹

¹Новосибирский государственный технический университет, проспект Карла Маркса 20, Новосибирск, Россия, 630073

Аннотация

Работа посвящена вопросам классификации данных ЭЭГ с использованием библиотеки алгоритмов машинного обучения Weka-3-8-4. Исследовано 34 алгоритма классификации, которые могут быть применены для решения задачи идентификации состояния глаз. Для проведения эксперимента использованы набор данных «Eeg-eye-state» для системы Weka и наборы данных, сформированные из записей ЭЭГ. Для автоматизации формирования наборов данных из записей разработан набор скриптов на языке Python.

Ключевые слова

ЭЭГ, Weka, машинное обучение, алгоритмы классификации

1. Введение

Задачи распознавания паттернов данных электроэнцефалографии (ЭЭГ) головного мозга на основе машинного обучения представляют собой быстрорастущее поле исследований. Цель исследования – провести анализ результатов работы алгоритмов машинного обучения, предоставляемых библиотекой WEKA-3-8-4, которые могут быть использованы для решения задачи идентификации состояния глаз. Методологической базой исследования являются работы [1] и [2], где исследуются вопросы определения двух состояния глаз (открытые или закрытые) по данным ЭЭГ. Работа организована следующим образом. В разделе 1 дается постановка задачи. В разделе 2 описываются материалы и методы исследования. В разделе 3 представляются результаты эксперимента. В разделе 4 подводятся итоги работы.

2. Материалы и методы

Для обучения классифицирующих моделей использована библиотека алгоритмов машинного обучения WEKA-3-8-4. Входные данные для инструмента машинного обучения Weka предоставляются в формате ARFF (Attribute-Relation File Format).

В работе использованы набор данных «Eeg-eye-state», размещенный в базе данных репозитория машинного обучения UCI, и набор данных «EEG Motor Movement/Imagery», размещенный в хранилище данных PhysioNet. Первый набор данных представляет собой файл в формате ARFF, в котором сохранены измерения с 14 датчиков нейрогарнитуры Emotiv EEG продолжительностью 117 секунд [3]. Второй набор состоит из одно- и двухминутных записей ЭЭГ, полученных от 109 добровольцев. ЭЭГ записаны с помощью 64-канального электроэнцефалографа и сохранены в формате EDF [4]. Из набора данных «EEG Motor Movement/Imagery» были взяты данные с 21 по 51 секунды для первых 10 субъектов для записей ЭЭГ с открытыми и закрытыми глазами в состоянии покоя. Для анализа были выбраны сигналы от 14 электродов, соответствующие схеме нейрогарнитуры Emotiv EPOC. Записи ЭЭГ были отфильтрованы в диапазоне альфа-полосы. Произведено удаление экземпляров с выбросами (пороговое значение: более чем 3 стандартных отклонения от медианы). Для автоматизации формирования наборов данных в формате ARFF из записей ЭЭГ был разработан набор скриптов на языке Python.

На первом этапе эксперимента на наборе данных «Eeg-eye-state» (набор 1) и на сформированном наборе данных для второго субъекта из «EEG Motor Movement/Imagery» (набор 2) выполнено обучение стандартных классификаторов WEKA, которые могут быть

