

УДК 004.891.3

ИНТЕРПРЕТАЦИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ДИАГНОСТИКИ ПСИХИЧЕСКИХ ЗАБОЛЕВАНИЙ ПО МРТ ГОЛОВНОГО МОЗГА

Петрова Д.А. (Университет ИТМО)

Научный руководитель – к.ф.м.н., доцент Михайлова Е.Г.
(Университет ИТМО)

В работе рассмотрены методы интерпретации сверточных нейронных сетей в задаче диагностики психических заболеваний по снимкам магнитно-резонансной томографии головного мозга. Описаны результаты тестирования алгоритмов Grad-CAM, Guided Backpropagation, Guided Grad-CAM и Meaningful Perturbation.

Введение. Одной из наиболее сложных проблем разработки систем поддержки принятия решений в медицине на основе моделей машинного обучения является обеспечение интерпретируемости результатов работы алгоритмов. Необходимость интерпретации обусловлена невозможностью постановки врачом клинического диагноза исключительно по бинарному предсказанию модели при большой цене ошибки. Несмотря на то, что качество классификации заболеваний по медицинским снимкам может быть достаточно высоким, далеко не все исследования в этой области предлагают инструментарий, позволяющий выявлять важные регионы на изображениях при принятии решений о наличии или отсутствии заболевания.

Основная часть. Входные данные представляют собой трехмерные снимки головного мозга. Датасет включает в себя 261 изображение, из них 49 принадлежат пациентам с биполярным расстройством, 40 и 50 с СДВГ и шизофренией соответственно, 122 – контрольная выборка. Процесс работы со снимками состоит из нескольких этапов: предобработки, классификации и интерпретации. Обучение классификатора, его применение и интерпретация производится на кросс-валидации по пяти фолдам.

Цель предобработки заключается в выделении на снимке непосредственно головного мозга и его преобразование к заранее заданной стандартной форме.

В качестве модели машинного обучения используется архитектура VoxResNet, изначально разработанная для анализа МРТ. Для каждого из заболеваний отдельно обучается бинарный классификатор.

Выбранные методы интерпретации – Guided Backpropagation, Grad-CAM, Guided Grad-CAM и Meaningful Perturbation. Первые три метода предназначены для обработки результатов предсказаний исключительно нейронных сетей, последний также подходит для черных ящиков. Результат работы каждого из алгоритмов после бинаризации – трехмерная маска, выделяющая важные для вердикта классификатора области на том или ином изображении. Ввиду особенностей рассматриваемых заболеваний индивидуальные маски могут быть усреднены для выделения характерных областей с аномалиями.

Визуальная и численная оценка качества интерпретации проводится в несколько этапов:

1. Визуальное сравнение масок, полученных с помощью обученной нейронной сети и нейронной сети со случайными весами.
2. Визуальное сравнение масок, полученных с помощью обученной нейронной сети на корректном датасете, и нейронной сети, обученной на датасете со случайно перемешанными метками классов.
3. Вычисление метрик качества для нейронной сети, обученной на выделенных с помощью масок областях, и их сравнение с первоначальными показателями.

В качестве основной метрики качества в данной работе используется ROC AUC.

Лучше других по совокупности критериев себя показали методы Grad-CAM и Meaningful Perturbation. В процессе работы была также разработана модификация метода Meaningful

Perturbation, основная идея которой состоит в поиске единственной маски для всего набора снимков МРТ.

Программное обеспечение для тестирования методов разработано на Python 3 с использованием PyTorch.

Следующим шагом является расширение датасета и набора методов интерпретации и проведение тестирования с другими нейросетевыми архитектурами. Задача построения алгоритма определения шизофрении на данный момент является наиболее перспективной из рассмотренных для дальнейших исследований, поскольку качество предсказания класса и интерпретации в ней достаточно высоко.

Выводы. Полученные результаты позволяют делать выводы о возможности интерпретации нейронных сетей для диагностики психических заболеваний и использования выбранных алгоритмов в клинической практике. Наибольший интерес для дальнейших исследований представляет задача диагностики шизофрении, однако требуется проведение тестирования на большем объеме данных, а также с использованием других архитектур нейронных сетей.

Петрова Д.А. (автор)

Подпись

Михайлова Е.Г. (научный руководитель)

Подпись