

Прогнозирование депрессивных расстройств у пациентов с рассеянным склерозом с использованием методов машинного обучения

Введение. Рассеянный склероз (РС) – тяжелое аутоиммунное нейродегенеративное и демиелинизирующее заболевание центральной нервной системы, приводящее к потере трудоспособности и прогрессирующей инвалидизации у лиц молодого возраста. В последнее время большое внимание уделяется когнитивным и нейропсихиатрическим расстройствам в клинической картине РС. Среди нарушений психической сферы наиболее частыми являются депрессивные эпизоды, гипомания, паническое расстройство, обсессивно-компульсивное расстройство. Распространенность депрессии у пациентов с РС значительно выше, чем у здоровой популяции. Данные симптомы, по мнению самих пациентов, значительно влияют на повседневную и профессиональную активность и ухудшают их качество жизни. Тщательная оценка нейропсихологической сферы пациентов с РС включает проведение разнообразного тестирования и опроса пациентов, однако является достаточно трудоемкой, длительной и порой сложной в интерпретации. Учитывая вышеизложенное, раннее прогнозирование депрессивных расстройств у пациентов с РС может значительно улучшить диагностический и лечебный процесс. Для поиска предикторов и создания прогностических моделей можно использовать современные методы машинного обучения, которые позволят более точно оценивать и прогнозировать исходы, чем «классические» статистические методы.

Цель. Построить модель машинного обучения для прогнозирования наличия у пациентов с РС депрессивных симптомов на основе предполагаемых предикторов.

Материалы и методы. На базе ГУ «Минский научно-практический центр хирургии, трансплантологии и гематологии» был проведен ретроспективный анализ группы из 57 пациентов с РС. Средний возраст составил 36 [32; 43] лет, распределение по полу: 25 мужчин (43,9%) и 32 женщины (56,1%), структура по формам: ремитирующая форма – у 46 (80,7%) пациентов, вторично-прогрессирующая форма – у 11 (19,3%) пациентов. В качестве предполагаемых предикторов использовались форма РС, пол, возраст, образование, наличие группы инвалидности, продолжительность заболевания, применение плазмафереза, лечение препаратами, изменяющими течение РС (ПИТРС), связь со стрессом, результаты девятиколышкового теста (9-NPT) для правой и левой руки, результат теста 25-футовой ходьбы, результат по шкале расширенной оценки инвалидизации

(EDSS), результаты когнитивных тестов по Монреальской шкале оценки когнитивных функций (MoCA), по таблицам Шульте (эффективность работы, степень вработываемости, психическая устойчивость), процент выполнения по символично-цифровому тесту (SDMT), результаты по психическому и физическому компоненту опросника по оценке качества жизни (SF-36). В качестве вероятных исходов выступало наличие или отсутствие депрессивных симптомов, которое определялось по шкале депрессии Бека (Beck Depression Inventory, BDI). Результат 10 и более баллов указывал на наличие депрессивной симптоматики. Для построения модели машинного обучения использовался язык программирования Python и следующие классификационные алгоритмы: логистическая регрессия, метод k-ближайших соседей (K-Nearest Neighbors, KNN), метод опорных векторов (Support Vector Machine, SVM), «решающее дерево» (decision tree), и ансамблевые методы «случайного леса» (random forest), классификатора голосований (Decision Classifier), и адаптивного бустинга (AdaBoost).

Результаты. Предварительно общая выборка была разделена на тренировочную (80%) и тестовую (20%). В качестве независимых переменных использовались значения вероятных предикторов, в качестве зависимой бинарный класс – отсутствие или наличие депрессивного синдрома. Для оценки эффективности работы модели использовались такие метрики, как точность (precision), полнота (recall), f1 score и ROC AUC score.

Наиболее эффективной моделью машинного обучения оказалась ансамблевая модель с помощью классификатора голосований, включающая в себя как модель «случайных лесов», так и модель на основе метода опорных векторов. Данная прогностическая модель позволяет прогнозировать наличие или отсутствие депрессивных симптомов у пациентов с рассеянным склерозом, основываясь на вышеперечисленных предикторах, с практически 100% точностью. Кроме данной модели хорошие результаты работы показали классификационные модели

Метрики эффективности работы построенных моделей машинного обучения

Метод	Precision	Recall	F1 score	ROC AUC score
Логистическая регрессия	0,8	0,8	0,8	0,89
KNN	0,8	0,8	0,8	0,94
SVM	0,83	1	0,91	0,89
«Решающее дерево»	0,67	0,8	0,73	0,76
«Случайный лес»	0,83	1	0,91	1
Классификатор голосований («случайный лес» + SVM)	1	1	1	1
Адаптивный бустинг («случайный лес» + SVM)	0,67	0,8	0,72	0,76

на основе таких алгоритмов, как «случайный лес», KNN, SVM. Метрики эффективности работы всех использованных моделей на тестовых данных представлены в таблице.

Наиболее значимыми предикторами депрессивного расстройства у пациентов с РС оказались физический компонент SF-36 (15,1%), процент выполнения по SDMT (11,09%), продолжительность заболевания (8,99%), психический компонент SF-36 (8,01%), эффективность работы по таблицам Шульте (6,7%), результат теста МОСА (6,32%), возраст пациента (5,99%), тест 25-футовой ходьбы (5,43%), результат девятиколышкового теста для правой руки (5,15%).

Заключение. Методы машинного обучения могут использоваться для построения эффективных систем прогнозирования, а также для поиска и анализа предикторов. Лучшие показатели работы наблюдаются у ансамблевых методов, объединяющих несколько алгоритмов машинного обучения. Наиболее эффективная модель машинного обучения имеет высокую точность предсказаний, а среди предложенных предикторов наибольший вклад в ее работу оказывают физический компонент SF-36, процент выполнения SDMT, продолжительность заболевания, психический компонент SF-36, эффективность работы по таблицам Шульте, результат теста МОСА, возраст, тест 25-футовой ходьбы, результат 9-НРТ для правой руки. В перспективе данная предикторная система может быть внедрена в Web-интерфейс.

Международный научно-практический журнал

НЕВРОЛОГИЯ и нейрохирургия

2023, том 13, № 1. Приложение

Восточная
Европа

Neurology and Neurosurgery. Eastern Europe
International Scientific Journal

2023 Volume 13 Number 1 Supplement



Валентин Коласанте и Поль Марк в балете «Ночь подходит к концу».
Музыка Людвиг ван Бетховена, хореография Бенжамена Мильпье.
Сцена Парко-делла-Музыка, Рим, 2023.

Тезисы XXII Республиканской
научно-практической конференции
для молодых специалистов

«СОВРЕМЕННЫЕ ДОСТИЖЕНИЯ НЕВРОЛОГИИ И НЕЙРОХИРУРГИИ»

Бобруйск, 2 июня 2023 года

ISSN 2226-0838 (Print)

ISSN 2414-3588 (Online)



9 772226 083006



ПРОФЕССИОНАЛЬНЫЕ
ИЗДАНИЯ