

В.А. Евстафьева
**WEB-ПРИЛОЖЕНИЕ ДЛЯ ПРОГНОЗИРОВАНИЯ РИСКА
ДЕПРЕССИВНЫХ РАССТРОЙСТВ У ПАЦИЕНТОВ
С РАССЕЯННЫМ СКЛЕРОЗОМ**

Научный руководитель: ассист. К.В. Благодичная
Кафедра нервных и нейрохирургических болезней
Белорусский государственный медицинский университет, г. Минск

V.A. Evstafieva
**WEB APPLICATION FOR PREDICTING THE RISK OF DEPRESSION
IN PATIENTS WITH MULTIPLE SCLEROSIS**

Tutor: assistant K. V. Blagochinnaya
Department of Nervous and Neurosurgical Diseases
Belarusian State Medical University, Minsk

Резюме. В данном исследовании было проведено нейропсихологическое тестирование пациентов с рассеянным склерозом, по результатам которого было разработано Web-приложение, позволяющее определять риск депрессивных расстройств на основе предикторов с использованием модели машинного обучения. Прогнозы данного приложения имеют высокую точность и могут использоваться в практической деятельности врачами различных специальностей.

Ключевые слова: рассеянный склероз, депрессивное расстройство, предикторы, машинное обучение, Web-приложение.

Resume. In this study, a neuropsychological testing of patients with multiple sclerosis was conducted, based on the results of which a Web application was developed that allows prognosing the risk of depressive disorders based on predictors using a machine learning model. The forecasts of this application are highly accurate and can be used in practice by doctors of various specialties.

Keywords: multiple sclerosis, depression, predictors, machine learning, web application.

Актуальность. Рассеянный склероз (РС) – тяжелое хроническое аутоиммунное демиелинизирующее и нейродегенеративное поражение центральной нервной системы (ЦНС), приводящее к потере трудоспособности и прогрессирующей инвалидизации у лиц преимущественно молодого возраста [1]. В последнее время возрос научный и практический интерес к когнитивным и нейропсихологическим расстройствам (НПР) в клинической картине РС, что связано с тем, что сами пациенты ставят на первое место по влиянию на качество жизни именно эти симптомы, тогда как врачи чаще проявляют больше беспокойства по поводу нарушения моторных функций [2]. Наиболее частыми НПР у пациентов с РС являются депрессивные расстройства, которые возможно являются не сопутствующим заболеванием, а проявлением обострения заболевания и по некоторым данным связаны с наличием свежих очагов демиелинизации в ЦНС [3]. Распространенность депрессии у пациентов с РС в 3-5 раз выше, чем в здоровой популяции и составляет около 30,5% [4], а распространенность отдельных депрессивных симптомов еще выше. Тщательная оценка наличия НПР у пациентов с РС является достаточно трудоемкой, длительной и порой сложной для интерпретации [5]. Для упрощения диагностического и лечебного процесса могут использоваться методы прогнозирования риска депрессивных расстройств. На данный момент

технологии искусственного интеллекта широко используются для мониторинга течения РС, в том числе с использованием нейровизуализационных данных [6]. Кроме того, методы машинного обучения можно использовать и для задач прогнозирования. Машинное обучение позволяет не только создавать предикторные модели, но еще и определять наиболее значимые переменные, а затем встраивать данные модели в удобный Web-интерфейс.

Цель: внедрить классификационную модель машинного обучения для прогнозирования риска депрессивных расстройств у пациентов с РС в Web-приложение.

Задачи:

1. Построить различные классификационные модели машинного обучения на тренировочной выборке данных.
2. Определить метрики эффективности данных моделей на тестовой выборке данных.
3. Выделить наиболее эффективную классификационную модель машинного обучения и определить для нее наиболее важные предикторы.
4. Внедрить самую эффективную модель в интерфейс Web-приложения.

Материалы и методы. Для построения моделей машинного обучения использовались данные 57 пациентов с верифицированным по критериям McDonald от 2017 г. [7] диагнозом «РС», исследованных на базе ГУ «Минский научно-практический центр хирургии, трансплантологии и гематологии». Половой состав исследованных пациентов: 25 мужчин (43,9%) и 32 женщины (56,1%); средний возраст – 36 [32; 43] лет; длительность заболевания варьирует от одного месяца до 30 лет; ремиттирующая форма наблюдалась у 46 (80,7%) пациентов, вторично-прогрессирующая форма – у 11 (19,3%) пациентов.

Для оценки депрессивных симптомов у исследованных пациентов использовалась шкала депрессии Бека (Beck Depression Inventory, BDI) [8]. Данный опросник является простым и удобным в использовании, а также валидизированным для использования у пациентов с неврологическими заболеваниями, в том числе и при РС. Средний результат опроса по BDI у исследованных пациентов составил 7 [3; 12] баллов, что соответствует норме, однако депрессивные симптомы наблюдались у 22 пациентов из 57 (38,6%).

В качестве независимых переменных (вероятных предикторов) для построения классификационных моделей машинного обучения использовались следующие признаки: форма РС, пол, возраст, образование, наличие или отсутствие группы инвалидности, продолжительность заболевания, применение плазмафереза, лечение препаратами, изменяющими течение РС (ПИТРС), связь со стрессом, результаты девятиколышкового теста (9-Hole Peg Test, 9-НРТ) для правой и левой руки, результат теста 25-футовой ходьбы, результат по шкале расширенной оценки инвалидизации (Expanded Disability Status Scale, EDSS), результаты когнитивных тестов по Монреальской шкале оценки когнитивных функций (Montreal Cognitive Assessment, МОСА), по таблицам Шульте (эффективность работы, степень вработываемости, психическая устойчивость), процент выполнения по символно-цифровому тесту

(Symbol Digit Modalities Test, SDMT), результаты по психическому и физическому компоненту опросника по оценке качества жизни (The Short Form-36, SF-36). В качестве вероятного исхода (зависимой переменной) выступало наличие или отсутствие депрессивных симптомов, которое определялось по шкале BDI. Результат 10 и более баллов указывал на наличие депрессивной симптоматики, 9 и менее баллов – на отсутствие депрессивной симптоматики.

Предобработка данных и построение моделей машинного обучения проводились с помощью языка программирования Python. Для построения моделей машинного обучения использовались такие классификационные модели, как 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 и их обсуждение. В качестве независимых переменных использовались значения вероятных предикторов, в качестве зависимой переменной – класс наличия либо отсутствия депрессивных симптомов. Для валидации результатов изначальный объем данных был предварительно поделен на две части – тренировочную (80%, для обучения модели) и тестовую (20%, для валидации модели) выборки, которые статистически значимо друг от друга не отличались. Для оценки эффективности работы модели использовались такие метрики, как точность (precision), полнота (recall), F1 score и ROC AUC score. Наилучшую эффективность показывают модели, у которых данные метрики стремятся к 1. Результаты работы различных классификационных моделей машинного обучения представлены в таблице 1.

Табл. 1. 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

Наиболее эффективной моделью машинного обучения оказалась ансамблевая модель классификатора голосований, включающая в себя как модель “случайных лесов”, так и модель на основе метода опорных векторов. Данная 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 на основе таких алгоритмов, как “случайный лес”, KNN, SVM.

Наиболее важными предикторами депрессивного расстройства у пациентов с РС согласно наиболее эффективной классификационной модели 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 (5,99%), тест 25-футовой ходьбы (5,43%), 9-НРТ для правой руки (5,15%). Все предикторы и их важности представлены на рисунке 1.

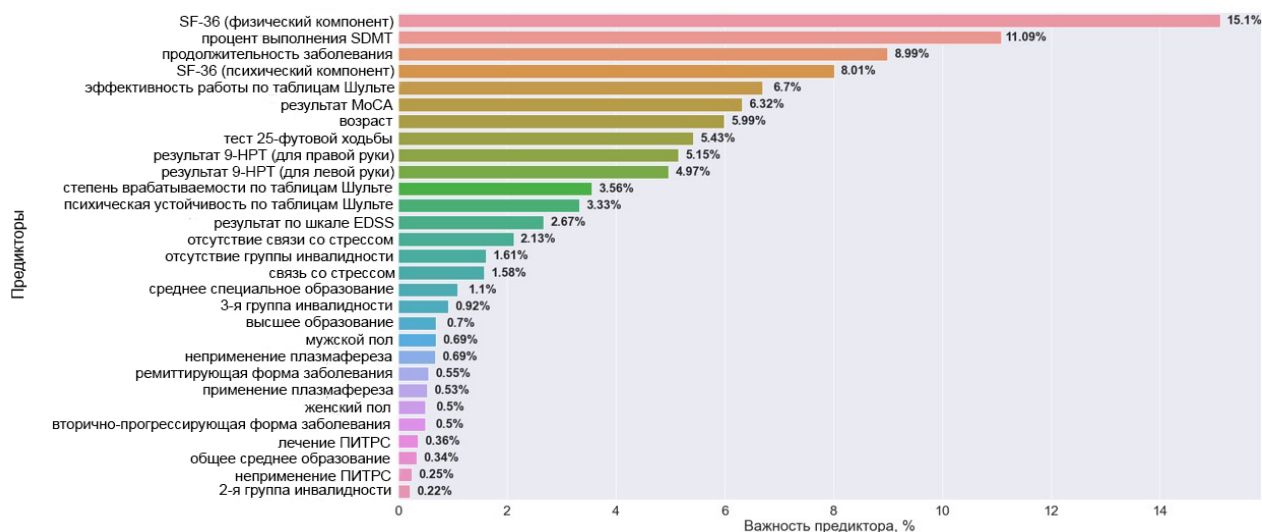


Рис. 1 – Важности предикторов наиболее эффективной классификационной модели машинного обучения

В дальнейшем самая эффективная классификационная модель была внедрена в интерфейс Web-приложения с использованием HTML, CSS и фреймворка Flask на языке программирования Python. Для работы с данным приложением необходимо заполнить все поля в форме на веб-странице и нажать на кнопку «Запустить прогноз», после чего в зависимости от результатов прогноза приложение выдаст текст «Риск депрессивных расстройств низкий» или «Риск депрессивных расстройств высокий». Работа приложения представлена на рисунках 2, 3.

Прогнозирование депрессии у пациентов с рассеянным склерозом

Пол:

Мужской

Женский

Возраст:

45

Форма заболевания:

Ремиттирующая форма

Вторично-прогрессирующая форма

Образование:

Среднее образование

Среднее специальное образование

Высшее образование

Рис. 2 – Интерфейс Web-приложения для прогнозирования депрессии у пациентов с РС

Введите отношение времени первой попытки к среднему времени прохождения (разделитель - точка)

Психическая устойчивость по таблицам Шульте:

Введите отношение времени четвертой попытки к среднему времени прохождения (разделитель - точка)

Психический компонент опросника SF-36:

Введите результат (разделитель - точка)

Физический компонент опросника SF-36:

Введите результат (разделитель - точка)

Запустить прогноз

Риск депрессивных расстройств низкий

Рис. 3 – Результат прогноза депрессии для одного из пациентов из исследованной группы

Выводы:

1. Наилучшие показатели метрики эффективности демонстрируют классификационные ансамблевые методы. Самая эффективная модель позволяет прогнозировать наличие или отсутствие депрессивных симптомов у пациентов с РС с практически 100% точностью.

2. Наиболее значимыми предикторами депрессивного расстройства у пациентов с РС для самой эффективной классификационной модели оказались физический компонент опросника 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%).

3. Данная классификационная модель была внедрена в Web-приложение, которое в последующем может эффективно использоваться в практической деятельности врачами различных специальностей вне зависимости от наличия знаний о программировании и машинном обучении.

Литература

1. Шмидт, Т.Е. Рассеянный склероз: от патогенеза через клинику к лечению / Т.Е. Шмидт, Н.Н. Яхно. – М.: МЕДпресс-информ, 2021. – 368 с.
2. Perceptions on the value of bodily functions in multiple sclerosis / C. Heesen [et al.] // *Acta Neurologica Scandinavica*. – 2018. – V. 137. – №. 3. – P. 356-362.
3. Prevalence of depression and anxiety in multiple sclerosis: a systematic review and meta-analysis / R.E. Boeschoten [et al.] // *Journal of the neurological sciences*. – 2017. – V. 372. – P. 331-341.
4. The link between multiple sclerosis and depression / A. Feinstein [et al.] // *Nature Reviews Neurology*. – 2014. – V. 10. – №. 9. – P. 507-517.
5. Patten, S.B. Depression in multiple sclerosis / S.B. Patten, R.A. Marrie, M.G. Carta // *International Review of Psychiatry*. – 2017. – V. 29. – №. 5. – P. 463-472.
6. Технологии искусственного интеллекта в мониторинге патоморфологических изменений центральной нервной системы при рассеянном склерозе / А.С. Федулов [и др.] // *Наука и инновации*. – 2023. – №. 2. – С. 75-83.
7. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria / A.J. Thompson [et al.] // *The Lancet Neurology*. – 2018. – V. 17. – №. 2. – P. 162-173.
8. Psychometric properties and validity of Beck Depression Inventory II in multiple sclerosis / R. Sacco [et al.] // *European journal of neurology*. – 2016. – V. 23. – №. 4. – P. 744-750.