

# ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ДИАГНОСТИКИ ПЕРВИЧНОЙ ОТКРЫТОУГОЛЬНОЙ ГЛАУКОМЫ ПО МУЛЬТИМОДАЛЬНЫМ ДАННЫМ

*Калиниченко А. В.*

*Научный руководитель: д-р мед. наук, доц. Качан Т. В.*

*Белорусский государственный медицинский университет, г. Минск*

**Резюме.** Для диагностики первичной открытоугольной глаукомы разработаны модели машинного обучения: Decision Tree, Logistic Regression, Random Forest, XGBoost (на основе анализа 31 глаза с диагнозом ПОУГ и 34 здоровых глаза). Наиболее эффективной моделью оказалась Decision Tree.

**Ключевые слова:** первичная открытоугольная глаукома модели машинного обучения.

**Актуальность.** Глаукома остается одной из наиболее инвалидизирующих заболеваний органа зрения. Согласно глобальным оценкам, около 6,9 миллиона человек в мире имеют умеренные и тяжелые нарушения зрения или слепоту, непосредственно вызванные глаукомой, значительную часть которых можно было бы предотвратить при своевременной диагностике и адекватном лечении. При этом подавляющее большинство этих случаев – примерно 70-90% – приходится на первичную открытоугольную глаукому (ПОУГ). [1].

В свете современных представлений глаукома позиционируется не как изолированная офтальмопатология, а как хроническое нейродегенеративное заболевание со специфическими механизмами апоптоза нейронов [2]. В основе развития глаукоматозной оптиконеуропатии лежит дегенерация ганглионарных клеток сетчатки (ГКС), сопровождающаяся атрофией их аксонов, которые формируют слой нервных волокон сетчатки

(СНВС) [3]. Это принципиально меняет подход к диагностике, смещая фокус с исключительно контроля внутриглазного давления (ВГД) на поиск ранних структурных и функциональных маркеров нейродегенерации. В данном контексте комплексный анализ взаимосвязи между толщиной ретинальных слоев (комплекса ГКС и СНВС), морфометрией диска зрительного нерва (ДЗН) и периметрическими параметрами с учетом системных факторов (таких как артериальная гипертензия и сахарный диабет) представляется крайне актуальным для углубления понимания патогенеза ПОУГ и улучшения ранней диагностики.

Использование методов искусственного интеллекта (ИИ) позволяет выявлять сложные, нелинейные закономерности данных, позволяя не только оценивать общие тенденции, но и строить прогностические модели для конкретного пациента. Таким образом, применение ИИ для анализа мультимодальных данных (ОКТ, пери-

метрия, системные показатели) является важным шагом в развитии диагностики и прогнозирования течения ПОУГ.

**Цель:** персонифицировать диагностику глаукомы на основе искусственного интеллекта.

**Задачи:**

1. Создать базу данных мультимодальных показателей пациентов с глаукомой и здоровых лиц.

2. Разработать и сравнить модели машинного обучения для диагностики ПОУГ.

3. Выявить ключевые признаки, влияющие на диагностику.

4. Оценить значимость лучшей модели в принятии клинических решений.

**Материалы и методы.** Данное ретроспективное исследование было проведено на базе УЗ «3-я городская клиническая больница имени Е.В. Клумова». В рамках работы сформирована база данных, включившая 65 глаз, которые были разделены на две группы: основную группу составил 31 глаз с диагнозом ПОУГ различных стадий, контрольную группу – 34 глаза здоровых пациентов. Критерии исключения для рассматриваемых групп включали наличие любой ретиальной патологии, а также миопию и гиперметропию высокой степени.

Для всех участников было выполнено комплексное офтальмологическое обследование, включающее: сбор анамнестических данных, измерение ВГД двумя методами (пневмотонометрия и тонометрия по Маклакову). Для оценки функционального состояния зрительной системы паци-

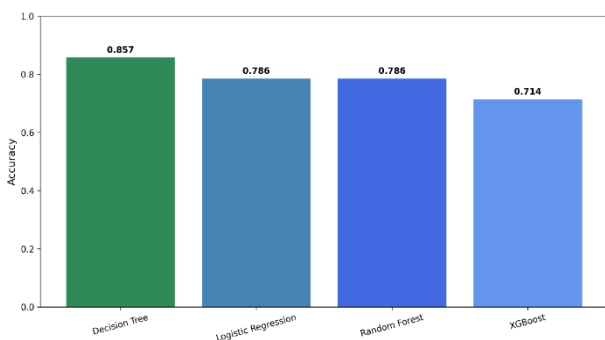
ентам проводилась кинетическая периметрия и автоматическая статическая компьютерная периметрия (MD, PSD) на анализаторе Zeiss Humphrey Visual Field Analyzer Model 720i (S/N 720i). Оценка структурных показателей сетчатки и ДЗН выполнялась с помощью оптической когерентной томографии на аппарате Topcon 3D OCT-1 Maestro с анализом следующих показателей: толщины сетчатки (общая толщина, толщина комплекса ганглиозных клеток и слоя нервных волокон), а также параметров ДЗН (площадь диска, площадь нейроретинального пояса, отношение площади экскавации к диску, горизонтальный и вертикальный диаметры). Качество полученных томограмм контролировалось по показателю силы сигнала, который составлял не менее 30 единиц.

Для создания диагностических моделей и их оценки были применены алгоритмы машинного обучения (МО) с использованием стандартных библиотек Python. Обработка данных выполнялась с помощью: pandas – для работы с табличными данными, numpy – для численных вычислений, scikit-learn и XGBoost – для машинного обучения, визуализация данных осуществлялась с помощью библиотек matplotlib и seaborn. Процесс построения модели включал этап предобработки данных с заменой пропущенных значений, кодированием категориальных признаков и масштабированием численных переменных. Все данные были разделены на трейновую и тестовую выборки в соотношении 80/20. После построения матрицы ошибок проводилась валидация модели по

метрикам: Accuracy, Precision, Recall, F1-score и ROC-AUC. Для интерпретации моделей выполнялся анализ значимости признаков (оценка Feature Importance). Статистический анализ включал описательную статистику с исследованием распределения переменных и выявлением выбросов.

**Результаты и их обсуждение.**

На основе полученных данных нами было создано несколько моделей МО: Decision Tree, Logistic Regression, Random Forest, XGBoost. Сравнительный анализ эффективности различных алгоритмов машинного обучения в диагностике ПОУГ показал, что модель дерева решений (Decision Tree) продемонстрировала наилучшую общую точность, достигшую 85.7% на тестовой выборке (рисунок 1). Критически важным результатом явилось то, что все протестированные модели показали чувствительность (Recall) на уровне 100%, что гарантирует полное выявление всех случаев глаукомы и отсутствие ложноотрицательных диагнозов.

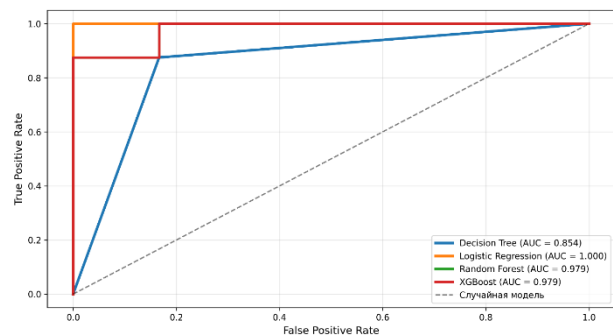


**Рис. 1** – Сравнение точности моделей МО

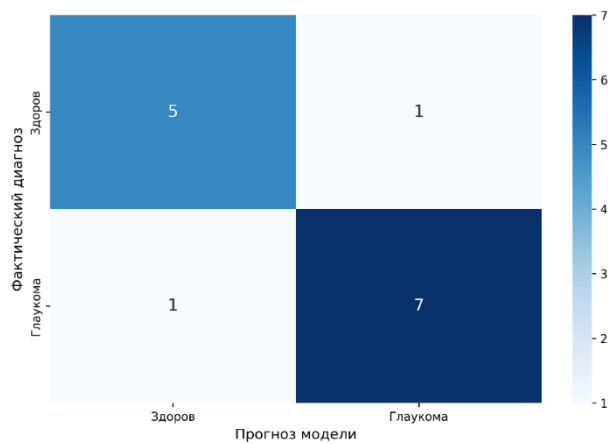
Анализ ROC-кривых подтвердил высокую дискриминативную способность моделей: Logistic Regression достигла площади под кривой (AUC) 1.000, Random Forest и XGBoost –

0.979, Decision Tree – 0,854 (рисунок 2).

Детальный разбор матрицы ошибок для оптимальной модели Decision Tree выявил, что все 8 случаев глаукомы были идентифицированы верно, однако наблюдалось 2 ложноположительных результата, что определяет специфичность модели на уровне 67% (рисунок 3).



**Рис. 2** – ROC-кривые моделей



**Рис. 3** – Матрица ошибок Decision Tree (лучшая модель)

Анализ значимости признаков (Feature Importance) позволил определить ключевые диагностические предикторы. Выявлены наиболее информативные показатели: кинетической периметрии (чувствительность в темпоральной зоне (важность = 0.806) и в назо-супериорной зоне (важность =

0.737)), а также показатели ОКТ (вертикальный диаметр ДЗН). Высокий вес также показали субъективное снижение остроты зрения (важность = 0.629) и наличие сахарного диабета (важность = 0.592). Данные результаты, где функциональные изменения полей зрения оказались наиболее информативными, полностью согласуются с современными представлениями о патогенезе ПОУГ как нейродегенеративного заболевания и подчеркивают ключевую роль комбинации функциональных нарушений и системных факторов риска.

#### **Выводы:**

1. В результате проведенного исследования были успешно разработаны и валидированы диагностические модели на основе алгоритмов Decision Tree, Logistic Regression, Random Forest, XGBoost.

2. Все протестированные модели машинного обучения показали максимальную чувствительность (100%), модель Decision Tree превзошла аналоги по специфичности (67%)

и общей точности (85.7%), что обуславливает её выбор в качестве оптимальной для клинического применения.

3. Методами машинного обучения выявлены и ранжированы ключевые диагностические предикторы ПОУГ. По суммарной оценке, наиболее информативными признаками являются: параметры кинетической периметрии (темпоральная и назо-супериорная зоны), показатели ОКТ (вертикальный диаметр ДЗН), субъективное снижение остроты зрения и наличие сахарного диабета.

4. Подтверждена клиническая релевантность и интерпретируемость разработанной модели Decision Tree, что создает основу для её практического внедрения. Для успешной интеграции в клиническую практику необходима дальнейшая валидация на независимой выборке большего объема, а также разработка пользовательского интерфейса, адаптированного для работы врачей-офтальмологов.

#### **Литература**

1. Всемирная организация здравоохранения. Всемирный доклад о проблемах зрения / Всемирная организация здравоохранения. – Женева: ВОЗ, 2020. – 180 с. – Лицензия: CC BY-NC-SA 3.0 IGO.
2. Almasieh, M. The molecular basis of retinal ganglion cell death in glaucoma / M. Almasieh, A. M. Wilson, B. Morquette [et al.] // Progress in Retinal and Eye Research. – 2012. – Vol. 31, № 2. – P. 152–181.
3. Качан, Т. В. Сравнительная характеристика оптической когерентной томографии и сканирующей лазерной поляриметрии в диагностике и мониторинге оптиконеуропатии у пациентов с глаукомой / Т. В. Качан, Л. Н. Марченко, Т. А. Бирич [и др.] // Офтальмология. Восточная Европа. – 2014. – № 4. – С. 186–194.

# APPLICATION OF MACHINE LEARNING METHODS FOR THE DIAGNOSIS OF PRIMARY OPEN-ANGLE GLAUCOMA BASED ON MULTIMODAL DATA

*Kalinichenko A. V.*

*Tutor: associate professor Kachan T. V.*

*Belarusian State Medical University, Minsk*

**Resume.** For the diagnosis of primary open-angle glaucoma, machine learning models have been developed: Decision Tree, Logistic Regression, Random Forest, XGBoost (based on the analysis of 31 eyes diagnosed with POAG and 34 healthy eyes). The Decision Tree turned out to be the most effective model.

**Keywords:** primary open-angle glaucoma machine learning models.