

УДК 004.89

Методы интеллектуального анализа данных в диагностике осложнений сахарного диабета

***О.С. Кротова¹, А.В. Ефанов¹, А.В. Кульшин¹,
А.Е. Жаксылыкова²***

*¹АлтГУ, г. Барнаул, ²ВКГУ им. С. Аманжолова,
г. Усть-Каменогорск, Казахстан*

Сахарный диабет – опасное хроническое заболевание, часто сопровождающееся серьезными осложнениями. Одним из наиболее распространенных осложнений сахарного диабета у детей и подростков является диабетическая полинейропатия, которая характеризуется поражением периферических нервных волокон, приводит к ухудшению качества жизни, и нередко становится причиной инвалидности [1, 2].

Целью исследования является применение методов интеллектуального анализа данных для автоматизации процесса диагностики диабетической полинейропатии.

Методы и подходы интеллектуального анализа данных являются мощным инструментом, который может применяться для решения различных прикладных задач, в том числе и задач медицинской диагностики. Методы интеллектуального анализа данных позволяют обнаружить скрытые закономерности в развитии заболеваний и их осложнений, создать модели автоматизированной диагностики, обученные на огромном количестве данных, сочетающем в себе информацию о различных проявлениях заболевания, формах его течения, способах диагностики.

Для реализации всех этапов анализа данных выбран высокоуровневый язык программирования Python [5].

Информационная база исследования представлена 3244 обезличенными медицинскими выписками из историй болезни детей и подростков Алтайского края, страдающих сахарным диабетом 1 типа.

Процесс анализа данных делится на несколько этапов:

- 1) Формирование базы данных;
- 2) Очистка данных;
- 3) Отбор признаков;
- 4) Построение и оптимизация моделей.

База данных состоит из 15 таблиц и содержит 96 показателей по каждому наблюдению. Для разработки моделей диагностики

диабетической полинейропатии сформирована выборка данных, в которую вошли 49 показателей: пол, возраст, длительность заболевания, показатели общего и биохимического анализа крови, общего анализа мочи. Выводы о наличии либо отсутствии диабетической полинейропатии сформированы исходя из результатов электронейромиографии верхних и нижних конечностей.

Первым этапом обработки данных является их очистка. В результате очистки данных найдены, обработаны и устранены ошибки ввода данных, нечисловые значения, дублирующиеся наблюдения.

Отбор признаков обусловлен необходимостью упростить модель, сократить время обучения и избежать переобучения модели. Классическими способами отбора признаков являются статистический анализ и применение алгоритмов автоматического отбора. Для реализации автоматического отбора признаков выбран метод последовательного обратного отбора. Задача диагностики диабетической полинейропатии, как и многие другие задачи медицинской диагностики, является нелинейной. Последовательный обратный отбор осуществлен с использованием двух классификаторов – метода опорных векторов с ядром из функции радиального базиса и метода k ближайших соседей. Для отбора признаков использовались 1100 наблюдений, в 500 из которых зафиксировано наличие полинейропатии. С помощью статистического анализа данных и автоматического отбора признаков выбраны 12 наиболее информативных признаков.

Для решения задачи автоматизации диагностики диабетической полинейропатии у детей и подростков выбраны следующие модели: метод опорных векторов, случайные лес, метод k ближайших соседей.

Оценка качества классификации проведена с помощью таких метрик, как точность, полнота, F -мера, чувствительность и специфичность. Метод исследования считается, оптимальным, если он высоко специфичен и высоко чувствителен. Однако в реальности повышение чувствительности неизбежно сопровождается потерей специфичности и наоборот, повышение специфичности сопряжено со снижением чувствительности. Все модели показали хорошее качество классификации. Наилучшее качество классификации показал метод опорных векторов – специфичность модели составляет 62%, чувствительность 82%.

Построенные модели диагностики диабетической полинейропатии позволят проводить скрининг осложнения вне стационара и без использования нейрофизиологических способов исследования, а также могут применяться в качестве системы поддержки принятия врачебных решений в диагностически неясных случаях.

Библиографический список

1. Алимова И.Л. Диабетическая нейропатия у детей и подростков: нерешенные проблемы и новые возможности // Российский вестник перинатологии и педиатрии. 2016. № 3. С. 114–123.
2. Дедов И.И., Кураева Т.Л., Петеркова В.А., Щербачёва А.Н. Сахарный диабет у детей и подростков. – М.: Универсум Паблишинг, 2002. – 392 с.
3. Дэви С., Арно М., Али М. Data Science и Big Data. Python и наука о данных. – СПб.: Питер, 2017. – 336 с.
4. Jelinek H.F., Cornforth D.J., Kelarev A.V. Machine learning methods for automated detection of severe diabetic neuropathy // Journal of Diabetic Complications & Medicine. 2016. Vol. 1. Issue 2. P. 1–7.
5. The Python Tutorial [Электронный ресурс]. – Режим доступа: <https://docs.python.org/3/tutorial/> – Загл. с экрана (Дата обращения 10.03.2020).

УДК 519.6

Вычисление интегральных топографических характеристик для многозональных снимков поверхности Земли в среде Matlab

*М.В. Куркина, В.А. Самарин, О.В. Самарина,
В.В. Славский
ЮГУ, г. Ханты-Мансийск*

Разработка автоматизированных информационных систем обработки многозональных снимков поверхности Земли является актуальной задачей в области обработки цифровых сигналов. Одними из наиболее эффективных являются алгоритмы, основанные на выделении контуров изображения. Контурный анализ содержит полную информацию о форме и контурах цифровых изображений, что позволяет эффективно использовать его при решении различных задач обработки цифрового изображения. В работе представлено описание практического метода вычисления интегральных топологических характеристик снимков поверхности Земли в среде Matlab.

Данные многозональной съемки в цифровом виде можно рассматривать как многомерную матрицу, где каждая двумерная матрица $u(x, y) \in [0, 255]$ является снимком земной поверхности в определенной зоне электромагнитного спектра.