

THE APPLICATION OF NEURO-FUZZY CASCADE NETWORK IN PREDICTION SYSTEMS OF PATIENTS WITH CHRONIC RENAL FAILURE

O. Yu. Kuznetsova¹, K. Mukapil²

¹Penza state university, Penza, Russia,

²Kazakh National Technical University named after K. I. Satpayev, Almaty, Kazakhstan.

E-mail: kaiyrkhan@mail.ru

Key words: chronic renal failure, cascading neuro-fuzzy network architecture of cascade neuro-fuzzy networks, fuzzy knowledge base.

Abstract. The application of neuro-fuzzy network with cascade architecture for constructing a prediction system of patients with chronic kidney disease is regarded. The network uses nodes with the ANFIS network and can be trained using optimization procedures. There has been made a study of the optimization algorithm of ANFIS networks using the method of sequential quadratic programming. The results of testing on the example of patients of the Penza Regional Clinical Hospital named after N. N. Burdenko showed a five time decrease of diagnostic errors.

УДК 004.032.26

ПРИМЕНЕНИЕ КАСКАДНОЙ НЕЙРО-НЕЧЕТКОЙ СЕТИ В СИСТЕМАХ ПРОГНОЗИРОВАНИЯ СОСТОЯНИЯ БОЛЬНЫХ С ХРОНИЧЕСКОЙ ПОЧЕЧНОЙ НЕДОСТАТОЧНОСТЬЮ

О.Ю. Кузнецова¹, К. Мукапил²

¹Пензенский государственный университет, Пенза, Россия

²Казахский национальный технический университет им. К. И. Сатпаева, Алматы, Казахстан

Ключевые слова: хроническая почечная недостаточность, каскадные нейро-нечеткие сети, архитектура каскадной нейро-нечеткой сети, нечеткая база знаний.

Аннотация. Рассматривается применение нейро-нечеткой сети с каскадной архитектурой для построения системы прогнозирования состояния больных с хронической почечной недостаточностью. Данная сеть использует узлы с сетью ANFIS и может быть обучена с помощью процедур оптимизации. Проведено исследование алгоритма оптимизации сети ANFIS с использованием метода последовательного квадратичного программирования. Результаты тестирования на примере пациентов Пензенской областной клинической больницы имени Н. Н. Бурденко показали уменьшение погрешности диагностики в пять раз.

Введение. Синдром эндогенной интоксикации (СЭИ) является одним из наиболее распространенных в клинической практике и характеризуется накоплением в тканях биологических продуктов, которые представляют собой результат реагирования на повреждающий фактор [1, 2]. Распространенным случаем СЭИ является хроническая почечная недостаточность (ХПН). По данным крупных популяционных регистров [1], таких как US Renal Data System и Российский регистр заместительной почечной терапии, распространенность хронической болезни почек (ХБП) составляет не менее 10%, достигая 20% и более у отдельных категорий лиц (пожилые, больные сахарным диабетом второго типа). Для сравнения: хроническая сердечная недостаточность

встречается у 1% населения, бронхиальная астма у 5% взрослого населения, сахарный диабет – у 4-10%. Признаки повреждения почек и снижение скорости клубочковой фильтрации выявляют, как минимум, у каждого десятого представителя общей популяции. При этом сопоставимые цифры были получены как в индустриальных странах с высоким уровнем жизни, так и в развивающихся странах со средним и низким доходом населения. Таким образом, задача ранней диагностики синдрома эндогенной интоксикации имеет большое значение. Задача ранней диагностики СЭИ затруднена сложностью использования специфических маркеров токсикоза. Для диагностики СЭИ существуют различные лабораторные методы, однако они являются дорогостоящими, в связи с чем встает задача создания менее дорогостоящих и более доступных методов диагностики [3]. Поэтому необходимо создавать менее дорогостоящие и более доступные методы диагностики.

Постановка задачи. Для обработки биомедицинских данных используют формальные методы, среди которых одними из наиболее популярных являются нейросетевые. Нейронные сети имеют возможность обучаться и обобщать накопленные знания и используются для задач классификации образов, распознавания, идентификации, прогнозирования, но не дают ответа на вопрос, как осуществляются эти процессы. Этот недостаток нейронных сетей решается в системах с нечетким выводом, в основе которых лежат понятия теории нечетких множеств и нечеткой логики [5].

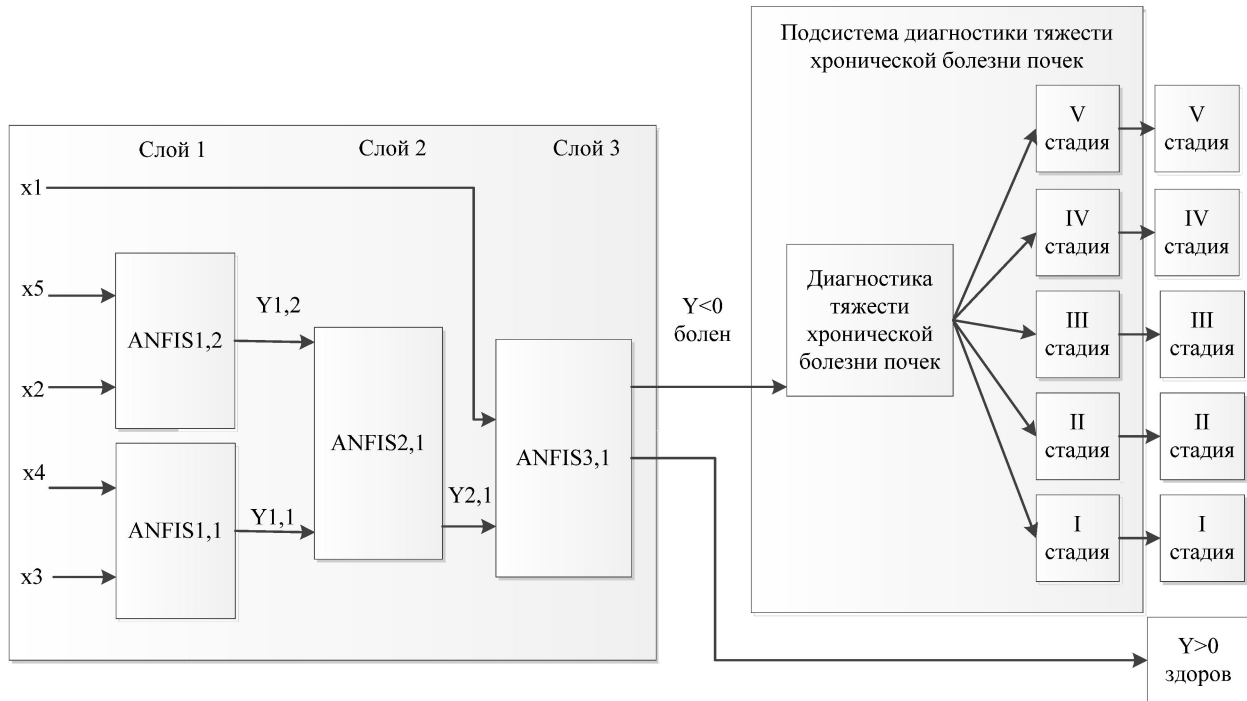
Системы с нечетким выводом позволяют объяснить получаемый с их помощью результат, позволяют закладывать в информационное поле априорный опыт врачей-экспертов, но они не имеют возможности обучаться и обобщать накопленные знания.

От перечисленных недостатков свободны системы на основе нейро-нечетких сетей. Выводы нейро-нечеткая сеть делает на основе базы знаний, в которой заключен априорный опыт эксперта, а параметры функций принадлежности настраиваются с использованием алгоритмов обучения нейронных сетей. В настоящей работе предпринимается попытка синтеза двух выше перечисленных сетей в каскадной нейро-нечеткой сети.

Разработка диагностической системы. Первоначально определяются входные переменные. Ими являются пять отобранных показателей – это общий белок, мочевины, креатинин, альбумины, билирубин. Единственным выходом сети является признак «больной-здоровый». Признак «здоровый» кодировался значением выхода $y = 1$, признак «больной» кодировался значением выхода $y = -1$. Таким образом, разрабатывалась модель нейро-нечеткой сети состояния пациента в зависимости от пяти показателей: x_1 – общий белок, x_2 – билирубин, x_3 – мочевины, x_4 – креатинин, x_5 – альбумины.

Структура адаптивной нейро-нечеткой сети (ANFIS) постоянна и зависит от количества входов. Так, для ANFIS с двумя входами имеем 28 настраиваемых коэффициентов, а с пятью – 232. Это ограничение для применения ANFIS с большим количеством входных параметров и малым количеством обучающих данных. Для сокращения числа настраиваемых коэффициентов разработана нейро-нечеткая сеть, имеющая каскадную архитектуру (рисунок). Данная сеть использует узлы с сетью ANFIS и может быть обучена с помощью процедур оптимизации, также сеть дополнена слоем, на котором определяется тяжесть ХБП.

В древовидной структуре в процессе поиска решения каждый узел дерева, начиная с листьев, помечается как решенная задача, если какая-нибудь подзадача либо все подзадачи решены. Процесс поиска продолжается до тех пор, пока не будет помечен корень дерева, т.е. решена исходная задача. Каждая вершина условия содержит некоторое высказывание, которое может принимать значения «Высокий уровень», «Средний уровень» или «Низкий уровень» входной переменной. Вершина вывода содержит одно или несколько предложений, описывающих некоторое промежуточное или окончательное заключение в виде набора показателей (соответствует выходам «здоровый» или «больной»). Корень дерева обозначается как вершина условия, содержащая высказывание, с которого начинается процесс логических рассуждений.



Структура каскадной нейро-нечеткой сети для диагностики СЭИ

Структура сети включает входной (нулевой) слой с внешними n -признаками и последующими k -слоями, каждый слой имеет m узлов. Количество слоев и узлов в каждом слое определяется по следующим формулам, пока не выполняются условия $t_k = 0$ и $m_k = 1$:

$$m_1 = \frac{m_0 - t_1}{d}, \quad m_2 = \frac{m_1 + t_1 - t_2}{d}, \quad m_3 = \frac{m_2 + t_2 - t_3}{d}, \quad \dots, \quad m_k = \frac{m_{k-1} + t_{k-1} - t_k}{d},$$

$$M = [m_1, m_2, \dots, m_k], \quad m_0 = l.$$

где k – количество слоев в диагностической системе; m_k – число, округленное до меньшего целого от деления (m_k/d) и определяющее количество узлов в k -слое; $t_k = m_{k-1} - dm_k$ – остаток от деления, определяющий количество незадействованных признаков в k -слое, переходящих на следующий слой.

Таким образом, создается сходящаяся каскадная структура, в которой выходы узлов слоя являются входами для следующего слоя. Узлом каскадной нейро-нечеткой сети является сеть ANFIS с заданным количеством входов d .

На каждый узел слоя должны последовательно подаваться группы диагностических признаков:

$$X_{i,j} = (x_1, x_2, \dots, x_d)_{i,j}, \quad i = \overline{1, 2, \dots, k}, \quad j = \overline{1, 2, \dots, m_i}$$

При формировании групп диагностических признаков используются следующие правила:

– признаки объединяются в группу, если они «не предсказывают» друг друга, т.е. когда они имеют наименьшие значения корреляции;

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

где x_i – значения, принимаемые в выборке X , y_i – значения, принимаемые в выборке Y , \bar{x} – средняя по X , \bar{y} – средняя по Y .

- каждый признак (диагностический) используется только в одном слое и только один раз;
- признаки (диагностические), которые не входят в группу в верхнем слое, подаются на нижние слои.

В качестве типового узла сети выступает сеть ANFIS $G_{i,j}(u)$ с набором правил. Ее параметры настраиваются при обучении известными методами оптимизации [6], соблюдая правила [7]:

- слои обучаются последовательно от первого к последнему;
- узлы в отдельном слое обучаются параллельно, причем на вход подается поставленная в соответствие группа, а на выход – вектор-столбец диагнозов.

Векторы-столбцы $g_{i,j}$, полученные при пересчете обученной сети $G_{i,j}(u)$ по заданной группе $(x_1, x_2, \dots, x_d)_{i,j}$, являются внутренними признаками, формирующими группы для ниже образующихся узлов. Выходом системы является выход сети последнего слоя.

Так как узлами каскадной нейро-нечеткой сети являются сети ANFIS, необходимо провести настройку их параметров. Особенностью сети ANFIS является то, что каждое правило нечетких продукций имеет постоянный весовой коэффициент равный 1. В данной работе параметры функций принадлежности входных переменных (показатели крови общего и биохимического анализов) подстраивать нельзя, так как это приведет к неверному диагнозу. Под обучением нейро-нечеткой сети понимается настройка параметров функций принадлежности и нечетких правил с помощью методов оптимизации. Особенностью обучения нейро-нечеткой сети ANFIS для диагностики синдрома эндогенной интоксикации на примере хронической почечной недостаточности является запрет изменения термов входных переменных «Общий белок», «Альбумины», «Мочевины» и «Креатинина», «Билирубин», что гарантирует правильную постановку диагноза.

Настройка сети ANFIS представляет собой нахождение весовых коэффициентов заключений правил. Процедура настройки по ошибке сети ANFIS на обучающей выборке модифицирует весовые коэффициенты. Тогда настройка сети ANFIS сводится к задаче оптимизации. Проведено исследование алгоритма оптимизации сети ANFIS с использованием метода последовательного квадратичного программирования (SQP – Sequential quadratic programming) с ограничениями. Настройка сети ANFIS рассматривается как задача параметрической оптимизации с ограничениями: требуется найти вектор w , обеспечивающий $f(w) \rightarrow \min$ при ограничениях $0 \leq g_i(w) \leq 1$ ($i = 1, 2, \dots, m$), где w – вектор оптимизируемых параметров (весовых коэффициентов правил), $f(w)$ – результат вывода по нечеткой базе знаний Сугено с параметрам (w) для входного вектора, $g_i(w)$ – некоторые скалярные функции векторного аргумента.

При выходе из подсистемы диагностики СЭИ с результатом «болен» происходит переход в подсистему диагностики тяжести хронической болезни почек (см. рисунок). Критерием тяжести является скорость клубочковой фильтрации (СКФ), которая рассчитывается по известным формулам СКД-EPI (Chronic Kidney Disease Epidemiology Collaboration), справедливым на любой стадии хронической болезни почек у представителей всех четырех рас [4].

Заключение. Завершив построение и обучение каскадной нейро-нечеткой сети, проведем оценку качества ее функционирования, вычисляя ошибки по каждому из ее узлов. Ошибкой сети будем считать неверно поставленный диагноз относительно «истинного» значения, т.е. известного врачебного диагноза. Исследования проводились на тех же выборках, что и ранее.

Тестирование каскадной нейро-нечеткой сети показало, что ошибка составила 1,7% на данных о больных с хронической почечной недостаточностью в терминальной стадии. Из них 0,7 % ошибка первого рода. Ошибку первого рода часто называют ложной тревогой, ложным срабатыванием или ложноположительным срабатыванием – например, анализ крови показал наличие заболевания, хотя на самом деле человек здоров.

Соответственно, ошибку второго рода иногда называют пропуском события или ложноотрицательным срабатыванием – человек болен, но анализ крови этого не показал.

Каскадная нечеткая нейронная сеть на данных о больных в ранней стадии хронической почечной недостаточности дает ошибку 11,8%, из них 1,8 % ошибка первого рода.

Таким образом, предложенная структура каскадной нейро-нечеткой сети позволила в пять раз уменьшить погрешность диагностики.

ЛИТЕРАТУРА

- [1] US Renal Data System, USRDS 2009 Annual Data Report: atlas of end-stage renal disease in the United States, National Institutes of Health, National Institutes of Diabetes and Digestive and Kidney Diseases, 2009.
- [2] Малахова М.Я. Метод регистрации эндогенной интоксикации: Пособие для врачей. – СПб.: Изд-во СПб МАПО, 1995. – 34 с.
- [3] Капустин Б.Б. Способы определения степени эндогенной интоксикации у больных абдоминальным сепсисом // Труды международного конгресса «Новые технологии в хирургии». – Ростов на Дону, 2005. – С. 47.
- [4] Levey A.S. CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A New Equation to Estimate Glomerular Filtration Rate / AS Levey, LA Stevens, CH Schmid, YL Zhang, AF Castro 3rd, HI Feldman, JW Kusek, P Eggers, F Van Lente, T Greene, J Coresh. Ann Intern Med. – 2009, 150:604-12.
- [5] Штовба С.Д. Введение в теорию нечетких множеств и нечеткую логику [Электронный ресурс] / С.Д. Штовба. – Режим доступа: <http://matlab.exponenta.ru/fuzzylogic/book1/>
- [6] Банди Б. Методы оптимизации. Вводный курс / Б. Банди; Под ред. В. А. Волынского. – М.: Радио и связь, 1988. – 128 с.
- [7] Безруков Н.С. Построение и моделирование адаптивной нейро-нечеткой системы в задачах медицинской диагностики / Н.С. Безруков, Е.Л. Еремин. – Информатика и системы управления. – 2005. – № 2(10). – С. 36-46.

REFERENCES

- [1] US Renal Data System, USRDS 2009 Annual Data Report: atlas of end-stage renal disease in the United States, National Institutes of Health, National Institutes of Diabetes and Digestive and Kidney Diseases, 2009.
- [2] Malakhova M.Ya. Method of registration of endogenous intoxication: A grant for doctors. SPb.: Publishing house of SPb of MAPO, 1995. 34 s.
- [3] Kapustin B.B. Ways of definition of degree of endogenous intoxication at patients with abdominal sepsis. Works of the international congress "New technologies in surgery". Rostov on Don, 2005. P. 47.
- [4] Levey A.S. CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A New Equation to Estimate Glomerular Filtration Rate. AS Levey, LA Stevens, CH Schmid, YL Zhang, AF Castro 3rd, HI Feldman, JW Kusek, P Eggers, F Van Lente, T Greene, J Coresh. Ann Intern Med. 2009. 150:604-12.
- [5] Shtovba S.D. Introduction to the theory of indistinct sets and fuzzy logic [An electronic resource]. S. D. Shtovba. Access mode: <http://matlab.exponenta.ru/fuzzylogic/book1/>
- [6] Bandi B. Optimization methods. Introduction course. B. Bandi; under the editorship of V. A. Volynsky. M.: Radio and communication, 1988. 128 p.
- [7] Bezrukov, N. S. Construction and modeling of adaptive neuro and indistinct system in problems of medical diagnostics. N. S. Bezrukov, E.L. Eremin. Informatics and control systems. 2005. N 2(10). P. 36-46.

БҮЙРЕК ЖҰМЫСЫНЫҢ ЖЕТКІЛІКСІЗ СОЗЫЛМАЛЫ АУЫРУЫН БОЛЖАУ ЖҮЙЕСІНДЕ КАСКАДТЫ АНЫҚ ЕМЕС - НЕЙРО ЖЕЛІСІН ҚОЛДАНУ

О. Ю. Кузнецова¹, К. Мукашп²

¹Пенза мемлекеттік университеті, Пенза, Ресей

²Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық университеті, Алматы, Қазақстан

Тірек сөздер: бүйрек жұмысының жеткіліксіз созылмалы ауыруы, каскадты анық емес – нейро желісі, каскадты анық емес - нейро желісінің архитектурасы.

Аннотация. Мақалада бүйрек жұмысының жеткіліксіз созылмалы ауыруымен ауыратын емделушілердің жағдайын болжау жүйесін құруда, каскадты анық емес – нейро желіні қолдану қарастырылған. Н. Н. Бурденко атындағы Пенза облыстық клиникалық ауруханасында өткізілген емделушілерді тестілеу нәтижесі каскадты анық емес – нейро желісін қолдану арқылы диагностикалық кемшіліктерді 5 есе азайтуға болатындығын көрсетті.

Поступила 15.01.2015 г.