

УДК 616.36+613.816

РАЗРАБОТКА МАТЕМАТИЧЕСКОЙ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ТЕЧЕНИЯ АЛКОГОЛЬНОЙ БОЛЕЗНИ ПЕЧЕНИ

© А.А. Арзамасцев, В.Б. Лифшиц, В.Н. Чичук

Ключевые слова: математическая модель; прогноз; алкогольная болезнь печени.

Математическое моделирование – быстро прогрессирующее научное направление. Актуальной проблемой современности является алкогольная болезнь печени (АБП). Ранее в мировой литературе не было предложено способов моделирования прогноза течения АБП. Несколько лет назад этот пробел был успешно восполнен с позиций кластерного и дискриминантного анализов на значительном числе наблюдений. В настоящей работе мы опять разработали математическую модель прогнозирования течения АБП другим способом, а именно методом искусственных нейронных сетей с учетом малого количества пациентов.

ВВЕДЕНИЕ

В настоящее время математическое моделирование стало одним из наиболее быстро прогрессирующих научных направлений, которое органично вошло во многие сферы научной и повседневной жизни [1–2]. На использовании математического моделирования как ядра интеллектуальных технологий базируются: построение экспертных систем различного назначения [1, 3–13], методики обработки научных данных и поиск скрытых закономерностей [7, 14–17], прогнозирование в различных сферах человеческой деятельности [10, 13, 18–20], управление социальными и техническими объектами и их системами [3, 9, 19, 21], идентификация внутренней структуры объектов различной природы [8, 19, 22–24], научно-технические расчеты [2, 12, 25].

Сущность методологии моделирования заключается в замене реального объекта его «образом» – моделью, изучение которой для исследователя оказывается более предпочтительным, нежели изучение самого объекта [2]. Такое предпочтение может быть вызвано следующими причинами [3]:

– исследование реального объекта в принципе невозможно. Такая ситуация имеет место, например, при прогнозировании в будущее или в прошлое, ведь машина времени пока еще не изобретена, и «посмотреть», что будет с объектом через некоторое время или что было с ним в прошлом, возможно только с помощью модели. Другой пример – проектирование новой технологии или технологического процесса; технология еще не существует в виде реальной конструкции, но на этапе ее проектирования уже необходимо просчитать основные технико-экономические и технологические параметры;

– исследование реального объекта в принципе возможно, но затруднено. Такая ситуация имеет место, например, когда проведение натуральных экспериментов является «дорогим удовольствием» или когда объект существует в единственном экземпляре, например, в процессе изучения природных объектов;

– реальный объект является чрезвычайно сложным, но исследователя интересует изучение поведения или свойств некоторой его относительно независимой части. В этой ситуации вместо сложного реального объекта возможно построить относительно простую модель, касающуюся лишь этой части; именно так и поступают в естественных науках, ограничиваясь в изучении явления лишь наиболее существенными параметрами, влияющими на его характеристики;

– моделируется гипотетический (реально не существующий) объект.

Построение математических моделей в различных областях знаний – естественных, технических науках, а также социальной сфере – принципиально возможно на основе теоретического, эмпирического и комбинированного подходов. В первом случае модель строится «за столом», т. е. ее основу составляют хорошо известные и проверенные теоретические положения и эмпирические данные. Никаких дополнительных экспериментов для ее разработки проводить не нужно. Эмпирический подход предполагает разработку модели исключительно на основе экспериментальных данных. Так обычно поступают в областях, в которых не существует развитой теоретической базы или трудно связать явления различных уровней. При таком подходе математическая модель обычно практически полностью лишена физичности, однако ее достоинствами являются относительная простота и универсальность используемых математических аппаратов. При комбинированном подходе математическая модель обычно строится исходя из теоретических соображений, а ее параметры определяются на основе эмпирических данных [3].

В нашем случае необходимо построение модели на основе эмпирического подхода, т. к. теоретические данные о связи выходного и входных параметров отсутствуют, но имеется значительный набор эмпирических данных по пациентам.

Способами реализации эмпирических моделей могут быть: регрессионный анализ, интерполирование, аппроксимация [15–17]; статистический анализ [8, 26];

корреляционный анализ совместно с аппроксимацией [15–17]; планированный эксперимент [15] и временные ряды [16]; искусственные нейронные сети [6, 8, 19, 27].

Для использования регрессионного анализа, интерполирования или аппроксимации входные и выходные переменные должны представлять собой действительные числа, что не выполняется в нашем случае (см. раздел 1). Для использования методов математической статистики и корреляционного анализа необходимо иметь значительное число измерений выходного параметра при одинаковых значениях входного вектора, что также не выполняется в нашем случае (см. раздел 1). Планированный эксперимент также неприменим ввиду того, что выборка эмпирических данных должна быть составлена таким образом, что каждый фактор должен принимать как минимум (для получения линейной модели) максимальное и минимальное значения. В нашем случае это потребовало бы $2^{24} = 16777216$ экспериментов! Временные ряды к данной ситуации неприменимы по своей сути.

В связи с вышеизложенным будем считать, что математическим аппаратом, приемлемым в нашем случае (малое число наблюдений), является аппарат искусственных нейронных сетей (ИНС) [6, 19, 27–28], ранее хорошо зарекомендовавший себя в качестве инструментария для построения моделей различных объектов на основе эмпирических данных [9, 16, 19–24, 29–35]. Необходимо отметить также, что имеется опыт использования этого аппарата для построения моделей в медицине [4, 29, 34, 36].

Достоинством моделей на ИНС (далее ИНС-моделей) является их достаточно простая интегрируемость в экспертные системы [4–5, 35, 37–38].

Алкогольное поражение печени и его осложнения остаются одной из самых частых причин смерти в Европе и США. Актуальным вопросом современности названа алкогольная болезнь печени, которая представляет собой последовательные формы стадии: стеатоз, хронический гепатит и цирроз печени [39]. Ранее в мировой литературе не было предложено способов моделирования прогноза течения алкогольной болезни печени [18]. Несколько лет назад этот пробел был успешно восполнен с позиций кластерного и дискриминантного анализов на значительном числе наблюдений (пациентов) [14, 26]. В настоящей работе мы опять представляем разработку математической модели прогнозирования течения алкогольной болезни печени другим способом, а именно методом ИНС на малом количестве наблюдений.

1. ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ ДАННЫХ НА ОСНОВЕ ИНС-МОДЕЛЕЙ

Нами были проанализированы данные здоровых людей, больных стеатозом, больных хроническим гепатитом, больных циррозом печени – всего 120 строк (по 30 строк в каждой группе). Данные представляют собой прямоугольную таблицу (матрицу), включающую 120 строк, каждая из которых соответствует пациенту, и 64 столбцов, каждый из которых в закодированном виде содержит данные об истории болезни, результатах инструментально-лабораторных обследований и других характеристиках пациентов.

После предварительного анализа данных из таблицы были исключены две строки и семь столбцов ввиду того, что информация, содержащаяся в них, не представлялась нам корректной. В окончательном виде таблица данных представляет собой прямоугольную матрицу, содержащую 118 строк и 58 столбцов. Каждая строка соответствует одному пациенту. Первые 57 столбцов таблицы содержат следующую информацию: x_1 – пол (1 – мужской, 2 – женский); x_2 – возраст (лет); x_3 – боль в правом подреберье (0 – нет, 1 – есть); x_4 – слабость, утомляемость (0 – нет, 1 – есть); x_5 – рвота (0 – нет, 1 – есть); x_6 – тошнота (0 – нет, 1 – есть); x_7 – аппетит (1 – снижен; 2 – сохранен; 3 – повышен); x_8 – стул (1 – нормальный, 2 – запор, 3 – понос); x_9 – метеоризм (0 – нет, 1 – есть); x_{10} – бессонница (0 – нет, 1 – есть); x_{11} – кожный зуд (0 – нет, 1 – есть); x_{12} – кровотечения (0 – нет, 1 – из десен, 2 – носовые, 3 – кожные геморрагии, 4 – маточные, 5 – геморроидальные); x_{13} – похудание (0 – нет, 1 – есть); x_{14} – количество ответов CAGE (0 – нет, 1 – 1 ответ, 2 – 2 ответа, 3 – 3 ответа, 4 – 4 ответа); x_{15} – доза приема этанола (мл) (1 – 40–80, 2 – 81–160, 3 – более 160); x_{16} – частота приема алкоголя (1 – менее одного раза в неделю, 2 – один раз в неделю, 3 – от двух до пяти раз в неделю, 4 – ежедневно); x_{17} – употребление алкоголя (лет); x_{18} – наследственность (0 – неотягощена, 1 – I степень родства, 2 – II ст., 3 – III ст., 4 – семью не знает); x_{19} – питание регулярное (0 – нет, 1 – да); x_{20} – соблюдение диеты (0 – нет, 1 – да); x_{21} – индекс массы тела (ИМТ); x_{22} – признаки «Сетки LeGo» (1 – менее 7, 2 – 7, 3 – более 7); x_{23} – кожа и слизистые (1 – норма, 2 – иктеричность склер, 3 – желтуха); x_{24} – сосудистые звездочки (0 – нет, 1 – да); x_{25} – печеночные ладони (0 – нет, 1 – да); x_{26} – боль при пальпации (0 – нет, 1 – да); x_{27} – край печени (1 – закругленный, 2 – острый); x_{28} – печень при пальпации (1 – норма, 2 – уменьшена, 3 – увеличена); x_{29} – консистенция печени (1 – мягкая, 2 – плотная); x_{30} – поверхность печени (1 – гладкая, 2 – бугристая); x_{31} – размеры по Курлову (1 – норма, 2 – уменьшены, 3 – увеличены); x_{32} – асцит (0 – нет, 1 – да); x_{33} – правая доля печени по УЗИ; x_{34} – левая доля печени по УЗИ; x_{35} – экзогенность печени (0 – пониженная, 1 – нормальная, 2 – повышенная); x_{36} – структура печени (1 – однородная, 2 – диффузно неоднородная); x_{37} – воротная вена; x_{38} – сосудистые коллатерали (0 – нет, 1 – да); x_{39} – сосудистый рисунок обеднен на периферии (0 – нет, 1 – да); x_{40} – длина селезенки; x_{41} – ширина селезенки; x_{42} – селезеночная вена; x_{43} – расширение вен пищевода (0 – нет, 1 – да); x_{44} – ЛПНП; x_{45} – ЛПВП; x_{46} – холестерин; x_{47} – триглицериды; x_{48} – билирубин общий; x_{49} – билирубин прямой; x_{50} – щелочная фосфатаза; x_{51} – АЛТ; x_{52} – АСТ; x_{53} – ГГТП; x_{54} – общий белок; x_{55} – альбумин; x_{56} – протромбиновое время; x_{57} – коэффициент де Ритиса. Последний 58 столбец содержит в закодированном виде сведения о болезни пациента: здоров – 1, стеатоз – 2, гепатит – 3, цирроз – 4.

Предварительный анализ данных проводили с помощью ИНС-модели, имеющей структуру, показанную на рис. 1.

Для обучения (параметрической идентификации) ИНС-модели использовали два метода нелинейного программирования – градиентный и Гаусса–Зейделя. Программы написаны на языке Borland Pascal. Далее приведены лучшие результаты обучения:

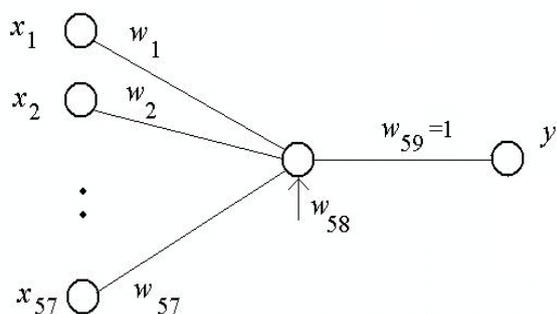


Рис. 1. Схема ИНС-модели исследуемого объекта

(c) ArzAA, 2012, The results of the ANN-model training.

Mean square error of the ANN-model is:
6,3794988741E-03

The weight coefficients of the model are:

- w[1] = -1,8547699920E-02
- w[2] = 1,9197999835E-03
- w[3] = 2,6866000358E-02
- w[4] = 5,9228200310E-02
- w[5] = 2,9152800110E-02
- w[6] = -4,6913000643E-02
- w[7] = -3,4535900241E-02
- w[8] = -1,3279000034E-02
- w[9] = -6,4870799951E-02
- w[10] = -6,2802500022E-02
- w[11] = -6,9975000558E-02
- w[12] = 3,4139800038E-02
- w[13] = -8,0389300197E-02
- w[14] = 1,7661499713E-02
- w[15] = 2,8744000154E-02
- w[16] = 6,5929994368E-04
- w[17] = -1,4520000160E-03
- w[18] = 1,1905299994E-02
- w[19] = -8,2615100432E-02
- w[20] = 2,7510700202E-02
- w[21] = 5,3518999946E-03
- w[22] = 5,4241500045E-02
- w[23] = 2,8163200097E-02
- w[24] = -1,2562830011E-01
- w[25] = -4,6490998617E-03
- w[26] = 1,4461599959E-02
- w[27] = 1,2737429930E-01
- w[28] = -2,9322800062E-02
- w[29] = -1,2927510040E-01
- w[30] = 2,0934340547E-01
- w[31] = 8,6009999728E-03
- w[32] = 5,5760800549E-02
- w[33] = 1,0510999883E-03
- w[34] = 1,3958999803E-03
- w[35] = 4,0980800074E-02
- w[36] = 4,9395930486E-01
- w[37] = -2,2709299828E-02
- w[38] = 5,6865571048E-01
- w[39] = 3,0519470729E-01
- w[40] = -8,2140001538E-04
- w[41] = -4,7016000231E-03
- w[42] = 3,3159994513E-04

- w[43] = 5,1734001029E-03
- w[44] = -1,3228590056E-01
- w[45] = -5,9964569308E-01
- w[46] = 1,3992470059E-01
- w[47] = 8,7720097421E-02
- w[48] = 2,2401999820E-03
- w[49] = -2,5986000139E-03
- w[50] = -8,9520001560E-04
- w[51] = 1,5108789795E-01
- w[52] = -3,5398099484E-02
- w[53] = 2,0129998411E-04
- w[54] = -2,8710002066E-04
- w[55] = 2,0114999856E-03
- w[56] = 1,9365299931E-02
- w[57] = 1,3803339449E-01
- w[58] = 7,1569881616E-01
- w[59] = 1,0000000000E+00
- Yexp = 1,00000 Ymod = 1,00753
- Yexp = 1,00000 Ymod = 0,93384
- Yexp = 1,00000 Ymod = 0,93977
- Yexp = 1,00000 Ymod = 0,99694
- Yexp = 1,00000 Ymod = 0,86882
- Yexp = 1,00000 Ymod = 1,05073
- Yexp = 1,00000 Ymod = 0,91645
- Yexp = 1,00000 Ymod = 1,03334
- Yexp = 1,00000 Ymod = 1,02998
- Yexp = 1,00000 Ymod = 0,95858
- Yexp = 1,00000 Ymod = 0,98839
- Yexp = 1,00000 Ymod = 0,99276
- Yexp = 1,00000 Ymod = 0,92581
- Yexp = 1,00000 Ymod = 0,98381
- Yexp = 1,00000 Ymod = 0,95288
- Yexp = 1,00000 Ymod = 0,90407
- Yexp = 1,00000 Ymod = 1,00230
- Yexp = 1,00000 Ymod = 1,21229
- Yexp = 1,00000 Ymod = 1,03254
- Yexp = 1,00000 Ymod = 1,15217
- Yexp = 1,00000 Ymod = 1,12515
- Yexp = 1,00000 Ymod = 1,01867
- Yexp = 1,00000 Ymod = 1,03112
- Yexp = 1,00000 Ymod = 0,95006
- Yexp = 1,00000 Ymod = 0,92558
- Yexp = 1,00000 Ymod = 1,08728
- Yexp = 1,00000 Ymod = 1,00568
- Yexp = 1,00000 Ymod = 1,04083
- Yexp = 1,00000 Ymod = 1,07091
- Yexp = 2,00000 Ymod = 2,04574
- Yexp = 2,00000 Ymod = 2,12966
- Yexp = 2,00000 Ymod = 2,06723
- Yexp = 2,00000 Ymod = 2,19269
- Yexp = 2,00000 Ymod = 1,90716
- Yexp = 2,00000 Ymod = 2,08194
- Yexp = 2,00000 Ymod = 2,01365
- Yexp = 2,00000 Ymod = 1,84804
- Yexp = 2,00000 Ymod = 1,92103
- Yexp = 2,00000 Ymod = 1,84240
- Yexp = 2,00000 Ymod = 1,92084
- Yexp = 2,00000 Ymod = 1,98969
- Yexp = 2,00000 Ymod = 2,03477
- Yexp = 2,00000 Ymod = 2,02708
- Yexp = 2,00000 Ymod = 1,97627
- Yexp = 2,00000 Ymod = 2,01613

$Y_{exp} = 2,00000$ $Y_{mod} = 2,18600$
 $Y_{exp} = 2,00000$ $Y_{mod} = 2,08577$
 $Y_{exp} = 2,00000$ $Y_{mod} = 2,10309$
 $Y_{exp} = 2,00000$ $Y_{mod} = 2,05112$
 $Y_{exp} = 2,00000$ $Y_{mod} = 1,94714$
 $Y_{exp} = 2,00000$ $Y_{mod} = 2,02190$
 $Y_{exp} = 2,00000$ $Y_{mod} = 1,98983$
 $Y_{exp} = 2,00000$ $Y_{mod} = 2,00851$
 $Y_{exp} = 2,00000$ $Y_{mod} = 1,87947$
 $Y_{exp} = 2,00000$ $Y_{mod} = 1,94626$
 $Y_{exp} = 2,00000$ $Y_{mod} = 2,06123$
 $Y_{exp} = 2,00000$ $Y_{mod} = 1,96342$
 $Y_{exp} = 2,00000$ $Y_{mod} = 1,96534$
 $Y_{exp} = 2,00000$ $Y_{mod} = 2,05407$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,07598$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,96080$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,10316$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,99076$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,01869$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,73791$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,12504$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,92759$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,90569$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,99537$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,86361$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,01355$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,09195$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,99385$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,14854$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,86655$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,92132$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,97535$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,02443$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,06293$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,99962$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,02850$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,96645$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,98589$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,92421$
 $Y_{exp} = 3,00000$ $Y_{mod} = 2,92616$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,01301$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,08325$
 $Y_{exp} = 3,00000$ $Y_{mod} = 3,05504$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,89836$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,90256$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,94302$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,02222$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,96367$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,94557$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,01030$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,02858$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,12914$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,07384$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,97775$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,09561$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,95890$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,03816$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,85164$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,15085$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,99167$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,95635$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,01921$

$Y_{exp} = 4,00000$ $Y_{mod} = 3,87839$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,01251$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,91081$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,01848$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,04106$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,07066$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,01380$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,97913$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,94718$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,99278$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,97690$

Здесь показаны значения весовых коэффициентов ИНС-модели и сравнение эмпирических данных (exp) и модельных расчетов (mod). Видно, что во всех случаях достигнуто хорошее совпадение при среднеквадратическом отклонении 0,0063794988741. На этом основании полученную ИНС-модель можно считать адекватной рассматриваемому объекту.

На рис. 2 показана относительная чувствительность входных каналов ИНС-модели. Эта зависимость позволяет судить о значимости того или иного параметра и его влиянии на конечный результат.

УДАЛЕНИЕ НЕЗНАЧИМЫХ ВХОДНЫХ ПАРАМЕТРОВ И КОРРЕКЦИЯ ИНС-МОДЕЛИ

В ходе дальнейшего анализа из рассмотрения были удалены все входные переменные, покрашенные на рис. 2 красным. Таким образом, в качестве входных переменных были оставлены лишь $x_1, x_{10}, x_{12}, x_{16}, x_{21}, x_{28}, x_{30}, x_{31}, x_{32}, x_{45}, x_{46}, x_{47}, x_{48}, x_{51}, x_{52}, x_{53}, x_{54}, x_{56}, x_{57}$ (всего 20 входных переменных, см. рис. 3).

При идентификации ИНС-модели получены следующие результаты:

(с) ArzAA, 2012. The results of the ANN-model training.

Mean square error of the ANN-model is: 4,2762614062E-02

The weight coefficients of the model are:

$w[1] = -1,2143769996E-01$
 $w[2] = 0,0000000000E+00$
 $w[3] = 0,0000000000E+00$
 $w[4] = 0,0000000000E+00$
 $w[5] = 0,0000000000E+00$
 $w[6] = 0,0000000000E+00$
 $w[7] = 0,0000000000E+00$
 $w[8] = 0,0000000000E+00$
 $w[9] = 0,0000000000E+00$
 $w[10] = -1,2157179958E-01$
 $w[11] = 1,8852800078E-02$
 $w[12] = 1,9906999362E-02$
 $w[13] = 0,0000000000E+00$
 $w[14] = 0,0000000000E+00$
 $w[15] = 0,0000000000E+00$
 $w[16] = 2,4407499972E-02$
 $w[17] = 0,0000000000E+00$
 $w[18] = 0,0000000000E+00$
 $w[19] = 0,0000000000E+00$
 $w[20] = 0,0000000000E+00$
 $w[21] = -2,0599984585E-04$
 $w[22] = 0,0000000000E+00$
 $w[23] = 0,0000000000E+00$

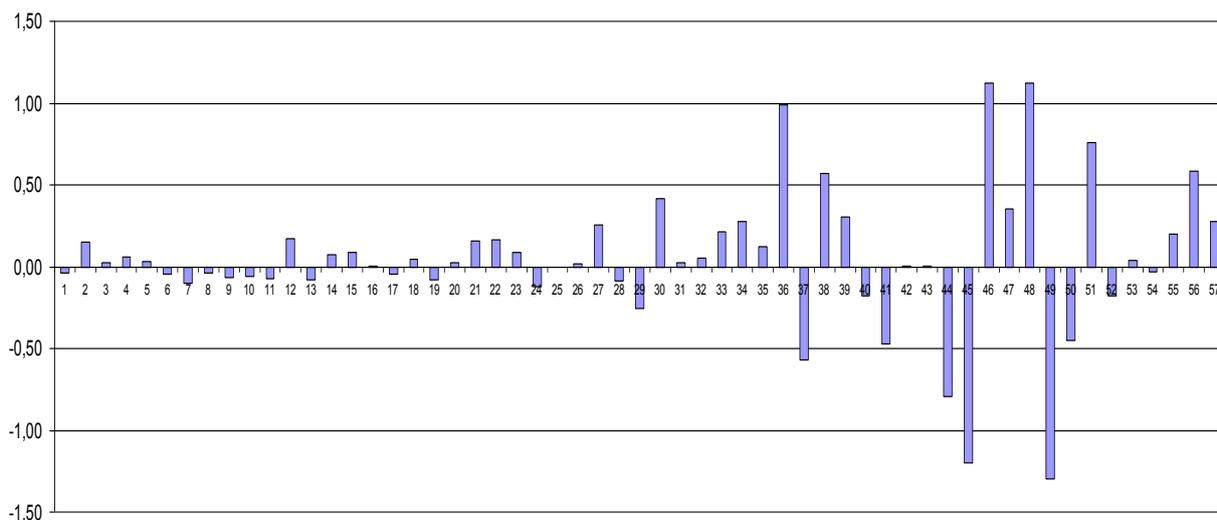


Рис. 2. Степень влияния различных факторов на заболевания печени при использовании шкалы: здоров – 1, стеатоз – 2, гепатит – 3, цирроз – 4 для группы из 118 пациентов. По оси абсцисс – факторы: x_1 – пол (1 – мужской, 2 – женский); x_2 – возраст (лет); x_3 – боль в правом подреберье (0 – нет, 1 – есть); x_4 – слабость, утомляемость (0 – нет, 1 – есть); x_5 – рвота (0 – нет, 1 – есть); x_6 – тошнота (0 – нет, 1 – есть); x_7 – аппетит (1 – снижен; 2 – сохранен; 3 – повышен); x_8 – стул (1 – нормальный; 2 – запор; 3 – понос); x_9 – метеоризм (0 – нет, 1 – есть); x_{10} – бессонница (0 – нет, 1 – есть); x_{11} – кожный зуд (0 – нет, 1 – есть); x_{12} – кровотечения (0 – нет, 1 – из десен, 2 – носовые, 3 – кожные геморрагии, 4 – маточные, 5 – геморроидальные); x_{13} – похудание (0 – нет, 1 – есть); x_{14} – количество ответов CAGE (0 – нет, 1 – 1 ответ, 2 – 2 ответа, 3 – 3 ответа, 4 – 4 ответа); x_{15} – доза приема этанола (мл) (1 – 40–80, 2 – 81–160, 3 – более 160); x_{16} – частота приема алкоголя (1 – менее одного раза в неделю, 2 – один раз в неделю, 3 – от двух до пяти раз в неделю, 4 – ежедневно); x_{17} – употребление алкоголя (лет); x_{18} – наследственность (0 – неотягощена, 1 – I степень родства, 2 – II степень, 3 – III степень, 4 – семья не знает); x_{19} – питание регулярное (0 – нет, 1 – да); x_{20} – соблюдение диеты (0 – нет, 1 – да); x_{21} – индекс массы тела (ИМТ); x_{22} – признаки «Сетки LeGo» (1 – менее 7, 2 – 7, 3 – более 7); x_{23} – кожа и слизистые (1 – норма, 2 – иктеричность склер, 3 – желтуха); x_{24} – сосудистые звездочки (0 – нет, 1 – да); x_{25} – печеночные ладони (0 – нет, 1 – да); x_{26} – боль при пальпации (0 – нет, 1 – да); x_{27} – край печени (1 – закругленный, 2 – острый); x_{28} – печень при пальпации (1 – норма, 2 – уменьшена, 3 – увеличена); x_{29} – консистенция печени (1 – мягкая, 2 – плотная); x_{30} – поверхность печени (1 – гладкая, 2 – бугристая); x_{31} – размеры по Курлову (1 – норма, 2 – уменьшены, 3 – увеличены); x_{32} – асцит (0 – нет, 1 – да); x_{33} – правая доля печени по УЗИ; x_{34} – левая доля печени по УЗИ; x_{35} – экзогенность печени (0 – пониженная, 1 – нормальная, 2 – повышенная); x_{36} – структура печени (1 – однородная, 2 – диффузно неоднородная); x_{37} – воротная вена; x_{38} – сосудистые коллатерали (0 – нет, 1 – да); x_{39} – сосудистый рисунок обеднен на периферии (0 – нет, 1 – да); x_{40} – длина селезенки; x_{41} – ширина селезенки; x_{42} – селезеночная вена; x_{43} – расширение вен пищевода (0 – нет, 1 – да); x_{44} – ЛПНП; x_{45} – ЛПВП; x_{46} – холестерин; x_{47} – триглицериды; x_{48} – билирубин общий; x_{49} – билирубин прямой; x_{50} – щелочная фосфатаза; x_{51} – АЛТ; x_{52} – АСТ; x_{53} – ГГТП; x_{54} – общий белок; x_{55} – альбумин; x_{56} – протромбиновое время; x_{57} – коэффициент де Ритиса. По оси ординат – относительная сила влияния факторов

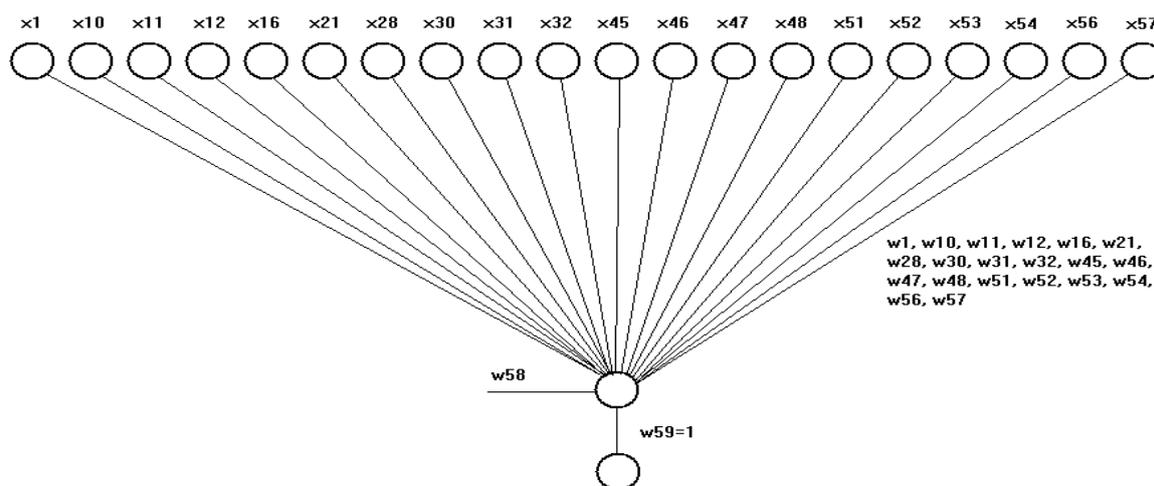


Рис. 3. Схема ИНС-модели объекта после удаления несущественных связей. Обозначения полностью соответствуют подрисункной подписи (рис. 2)

w[24] = 0,000000000E+00	Yexp = 1,00000 Ymod = 0,99077
w[25] = 0,000000000E+00	Yexp = 1,00000 Ymod = 1,04643
w[26] = 0,000000000E+00	Yexp = 1,00000 Ymod = 1,03055
w[27] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,89294
w[28] = -5,9784999510E-03	Yexp = 2,00000 Ymod = 2,40652
w[29] = 0,000000000E+00	Yexp = 2,00000 Ymod = 2,20114
w[30] = 4,0571699855E-02	Yexp = 2,00000 Ymod = 2,31178
w[31] = 4,6150900147E-02	Yexp = 2,00000 Ymod = 2,04966
w[32] = 5,8581599945E-02	Yexp = 2,00000 Ymod = 1,55990
w[33] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,57648
w[34] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,72033
w[35] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,65646
w[36] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,93965
w[37] = 0,000000000E+00	Yexp = 2,00000 Ymod = 2,01089
w[38] = 0,000000000E+00	Yexp = 2,00000 Ymod = 2,07124
w[39] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,93794
w[40] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,74820
w[41] = 0,000000000E+00	Yexp = 2,00000 Ymod = 2,09441
w[42] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,89442
w[43] = 0,000000000E+00	Yexp = 2,00000 Ymod = 2,29335
w[44] = 0,000000000E+00	Yexp = 2,00000 Ymod = 2,04918
w[45] = -5,7476529241E-01	Yexp = 2,00000 Ymod = 2,23761
w[46] = 1,8004859999E-01	Yexp = 2,00000 Ymod = 1,83325
w[47] = 6,7693839925E-01	Yexp = 2,00000 Ymod = 1,61348
w[48] = 2,0159994510E-04	Yexp = 2,00000 Ymod = 1,88056
w[49] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,86060
w[50] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,80515
w[51] = -1,5034569132E-01	Yexp = 2,00000 Ymod = 1,82070
w[52] = 1,4037469998E-01	Yexp = 2,00000 Ymod = 1,71084
w[53] = 7,8009743150E-04	Yexp = 2,00000 Ymod = 2,05074
w[54] = 3,2601999832E-03	Yexp = 2,00000 Ymod = 1,87284
w[55] = 0,000000000E+00	Yexp = 2,00000 Ymod = 1,80693
w[56] = 3,0584799999E-02	Yexp = 2,00000 Ymod = 1,86630
w[57] = -1,2703210199E-01	Yexp = 3,00000 Ymod = 3,00636
w[58] = -1,3022809961E-01	Yexp = 3,00000 Ymod = 3,23293
w[59] = 1,000000000E+00	Yexp = 3,00000 Ymod = 3,01874
Yexp = 1,00000 Ymod = 1,22103	Yexp = 3,00000 Ymod = 3,18674
Yexp = 1,00000 Ymod = 1,28429	Yexp = 3,00000 Ymod = 2,95776
Yexp = 1,00000 Ymod = 1,00829	Yexp = 3,00000 Ymod = 2,92724
Yexp = 1,00000 Ymod = 1,06473	Yexp = 3,00000 Ymod = 2,91934
Yexp = 1,00000 Ymod = 0,95554	Yexp = 3,00000 Ymod = 2,45790
Yexp = 1,00000 Ymod = 1,01253	Yexp = 3,00000 Ymod = 2,62119
Yexp = 1,00000 Ymod = 1,04146	Yexp = 3,00000 Ymod = 2,96377
Yexp = 1,00000 Ymod = 1,11902	Yexp = 3,00000 Ymod = 3,02690
Yexp = 1,00000 Ymod = 1,17056	Yexp = 3,00000 Ymod = 2,99351
Yexp = 1,00000 Ymod = 0,94688	Yexp = 3,00000 Ymod = 3,01267
Yexp = 1,00000 Ymod = 1,11420	Yexp = 3,00000 Ymod = 3,17221
Yexp = 1,00000 Ymod = 1,11852	Yexp = 3,00000 Ymod = 2,98506
Yexp = 1,00000 Ymod = 1,14888	Yexp = 3,00000 Ymod = 2,78911
Yexp = 1,00000 Ymod = 1,45291	Yexp = 3,00000 Ymod = 3,14924
Yexp = 1,00000 Ymod = 1,33313	Yexp = 3,00000 Ymod = 3,19147
Yexp = 1,00000 Ymod = 1,10546	Yexp = 3,00000 Ymod = 2,99678
Yexp = 1,00000 Ymod = 1,33453	Yexp = 3,00000 Ymod = 3,00045
Yexp = 1,00000 Ymod = 1,59261	Yexp = 3,00000 Ymod = 2,99625
Yexp = 1,00000 Ymod = 1,03731	Yexp = 3,00000 Ymod = 2,74893
Yexp = 1,00000 Ymod = 1,25177	Yexp = 3,00000 Ymod = 2,82964
Yexp = 1,00000 Ymod = 1,32009	Yexp = 3,00000 Ymod = 2,89268
Yexp = 1,00000 Ymod = 0,89745	Yexp = 3,00000 Ymod = 3,06192
Yexp = 1,00000 Ymod = 0,87656	Yexp = 3,00000 Ymod = 2,67590
Yexp = 1,00000 Ymod = 0,98375	Yexp = 3,00000 Ymod = 3,17853
Yexp = 1,00000 Ymod = 0,96199	Yexp = 3,00000 Ymod = 3,00302
Yexp = 1,00000 Ymod = 1,00642	Yexp = 3,00000 Ymod = 2,81939

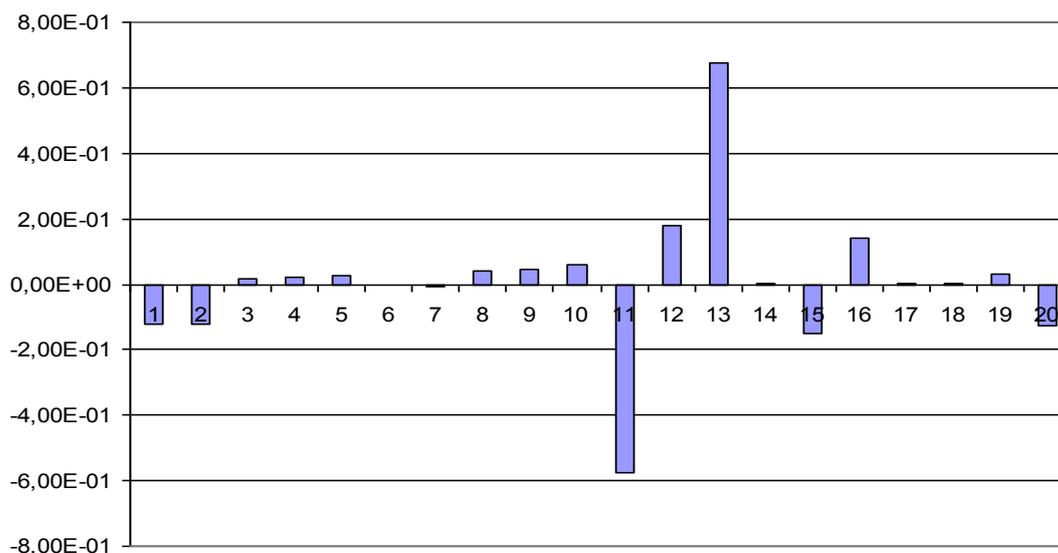


Рис. 4. Относительная значимость двадцати входных каналов: 1 – x_1 , 2 – x_{10} , 3 – x_{11} , 4 – x_{12} , 5 – x_{16} , 6 – x_{21} , 7 – x_{28} , 8 – x_{30} , 9 – x_{31} , 10 – x_{32} , 11 – x_{45} , 12 – x_{46} , 13 – x_{47} , 14 – x_{48} , 15 – x_{51} , 16 – x_{52} , 17 – x_{53} , 18 – x_{54} , 19 – x_{56} , 20 – x_{57} . Обозначения переменных соответствуют рис. 2

$Y_{exp} = 4,00000$ $Y_{mod} = 3,98143$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,26725$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,14265$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,21386$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,64652$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,13084$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,16154$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,47027$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,06086$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,18294$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,78632$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,12247$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,65098$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,72885$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,07308$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,22530$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,91503$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,97444$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,92214$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,10814$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,98791$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,88617$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,08049$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,77892$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,99144$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,69627$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,09765$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,61648$
 $Y_{exp} = 4,00000$ $Y_{mod} = 3,78305$
 $Y_{exp} = 4,00000$ $Y_{mod} = 4,19842$

Средняя квадратичная погрешность по сравнению с первой ИНС-моделью возросла до 0,0428, однако ее уровень остался приемлемым для того, чтобы считать модель адекватной.

Таким образом, нами была разработана адекватная ИНС-модель прогнозирования течения алкогольной болезни печени.

ЛИТЕРАТУРА

1. Анищенко В.С., Булдакова Т.И., Лифищ В.Б., Довгалецкий П.Я., Гриднев В.И. Концептуальная модель виртуального центра охраны здоровья населения // Информационные технологии. 2009. № 12. С. 54-59.
2. Самарский А.А., Михайлов А.П. Математическое моделирование. М.: Физматлит, 2002. 320 с.
3. Арзамасцев А.А. Математическое и компьютерное моделирование. Тамбов: Издат. дом ТГУ им. Г.Р. Державина, 2010. 257 с.
4. Арзамасцев А.А., Зенкова Н.А., Неудахин А.В. Технология построения медицинской экспертной системы на основе аппарата искусственных нейронных сетей // Информационные технологии. 2009. № 8. С. 60-63.
5. Арзамасцев А.А., Зенкова Н.А., Неудахин А.В. Формализация проблемы разработки экспертной информационной системы с развивающимся интеллектуальным ядром на базе ИНС-моделей // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2010. Т. 15. Вып. 1. С. 287-290.
6. Барский А.Б. Нейронные сети: распознавание, управление, принятие решений. М.: Финансы и статистика, 2004. 176 с.
7. Кубенский А.А. Структуры и алгоритмы обработки данных. СПб.: Изд-во БХВ-Петербург, 2004. 466 с.
8. Булдакова Т.И., Лифищ В.Б., Суятинов С.И., Колентьев С.В. Статистический и нейросетевой методы идентификации и прогнозирования в медицине // Информационные технологии. 2004. № 3. С. 67-71.
9. Зенкова Н.А. Компьютерное моделирование в психологии. Тамбов: ИМФИ ТГУ им. Г.Р. Державина, 2007. 55 с.
10. Лифищ В.Б., Булдакова Т.И., Екимова Н.В., Игнатьева Е.В., Суятинов С.И. Прогнозирование развития холестерина желчного пузыря // Технологии живых систем. 2009. Т. 6. № 2. С. 52-59.
11. Лифищ В.Б., Сернов С.П. Математическое моделирование в диагностике АБП // Вестник современной клинической медицины. 2010. Т. 3. Приложение 1. С. 106-107.
12. Мышкис А.Д. Элементы теории математических моделей. М.: Кн. дом «ЛИБРОКОМ», 2009. 192 с.
13. Сушкова Н.В., Лифищ В.Б. Прогнозирование развития заболевания желчного пузыря. Значение клинико-сонографических и биохимических критериев. Saarbrücken: LAP LAMBERT Academic Publishing, 2011. 149 p.
14. Булдакова Т.И., Лифищ В.Б., Сернов С.П., Суятинов С.И. Проблема диагностики алкогольной болезни печени с позиций кластерного анализа // Вестник новых медицинских технологий. 2010. Т. 17. № 1. С. 55-57.
15. Дюк В. Обработка данных на ПК в примерах. Санкт-Петербург: Москва; Харьков; Минск, 1997. 231 с.
16. Торин Ю.Н., Макаров А.А. Анализ данных на компьютере. М.: ИНФА-М, Финансы и статистика, 2003. 544 с.

17. *Тюрин Ю.Н., Макаров А.А.* Статистический анализ данных на компьютере. М.: ИНФА-М, 1998. 528 с.
18. *Сернов С.П., Лифшиц В.Б., Скворцов Ю.И., Субботина В.Г., Мартынова А.Г., Сулковская Л.С., Шульгин В.И., Сучилина Л.А.* Актуальные проблемы прогнозирования алкогольной болезни печени // Саратовский научно-медицинский журнал. 2010. Т. 6. № 1. С. 94-100.
19. *Арзамасцев А.А., Зенкова Н.А.* Искусственный интеллект и распознавание образов. Тамбов: Издат. дом ТГУ им. Г.Р. Державина, 2010. 196 с.
20. *Kryuchin O.V., Arzamastsev A.A., Zenkova N.A., Troitzsch K.G., Sletkov D.V.* Simulating medical objects simulation using an artificial neural network whose structure is based on adaptive resonance theory // Institut für Wirtschafts- und Verwaltungsinformatik Fachbereich Informatik Universität Koblenz-Landau. № 14/2011. URL: http://www.uni-koblenz.de/~fb4reports/2011/2011_14_Arbeitsberichte.pdf. Загл. с экрана.
21. *Арзамасцев А.А., Зенкова Н.А.* Моделирование в психологии на основе аппарата искусственных нейронных сетей. Тамбов: ИМФИ ТГУ им. Г.Р. Державина, 2003. 106 с.
22. *Арзамасцев А.А., Зенкова Н.А.* Использование аппарата искусственных нейронных сетей для разработки систем психологического и профессионального тестирования // Математические методы и информационные технологии в экономике, социологии и образовании: сборник статей 11 Международной научно-технической конференции. Пенза, 2003. С. 258-260.
23. *Арзамасцев А.А., Зенкова Н.А.* Нейросетевая технология психологического тестирования степени готовности абитуриентов к образовательной деятельности // Новые биокибернетические и телемедицинские технологии 21 века для диагностики заболеваний человека (НБИГТ-21): материалы междисциплинарной конференции с международным участием. Петрозаводск, 2002. С. 37.
24. *Арзамасцев А.А., Зенкова Н.А.* Система психологического тестирования на основе аппарата искусственных нейронных сетей // Искусственный интеллект. 2004. № 2. С. 237-242.
25. *Тарасевич Ю.Ю.* Математическое и компьютерное моделирование. М.: Эдиториал УРСС, 2004. 152 с.
26. *Лифшиц В.Б., Булдакова Т.И., Иенатьева Е.В., Сернов С.П.* Скрининг групп риска алкогольной болезни печени среди населения // Системный анализ и управление в биомедицинских системах. 2010. Т. 9. № 2. С. 370-375.
27. *Осовский С.* Нейронные сети для обработки информации. М.: Финансы и статистика, 2004. 344 с.
28. *Уоссермен Ф.* Нейрокомпьютерная техника: теория и практика. М.: Мир, 1992. 240 с.
29. *Крючин О.В., Зенкова Н.А.* Использование искусственных нейронных сетей для решения задач классификации на примере моделирования медицинского объекта // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2011. Т. 16. Вып. 3. С. 789-792.
30. *Arzamastsev A.A., Zenkova N.A., Troitzsch K.G., Neuakhin A.V.* Technology of intellectual information system design for estimation of social objects // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2011. Т. 16. Вып. 1. С. 169-173.
31. *Kryuchin O.V., Arzamastsev A.A., Troitzsch K.G.* A parallel algorithm for selecting activation functions of an artificial network // Institut für Wirtschafts- und Verwaltungsinformatik Fachbereich Informatik Universität Koblenz-Landau. № 12/2011. URL: http://www.uni-koblenz.de/~fb4reports/2011/2011_12_Arbeitsberichte.pdf. Загл. с экрана.
32. *Kryuchin O.V., Arzamastsev A.A., Troitzsch K.G.* A universal simulator based on artificial neural networks for computer clusters // Institut für Wirtschafts- und Verwaltungsinformatik Fachbereich Informatik Universität Koblenz-Landau. № 2/2011. URL: http://www.uni-koblenz.de/~fb4reports/2011/2011_02_Arbeitsberichte.pdf. Загл. с экрана.
33. *Kryuchin O.V., Arzamastsev A.A., Troitzsch K.G.* Comparing the efficiency of serial and parallel algorithms for training artificial neural networks using computer clusters // Institut für Wirtschafts- und Verwaltungsinformatik Fachbereich Informatik Universität Koblenz-Landau № 13/2011. URL: http://www.uni-koblenz.de/~fb4reports/2011/2011_13_Arbeitsberichte.pdf. Загл. с экрана.
34. *Kryuchin O.V., Arzamastsev A.A., Troitzsch K.G.* The prediction of currency exchange rates using artificial neural networks // Institut für Wirtschafts- und Verwaltungsinformatik Fachbereich Informatik Universität Koblenz-Landau. № 4/2011. http://www.uni-koblenz.de/~fb4reports/2011/2011_04_Arbeitsberichte.pdf. Загл. с экрана.
35. *Zenkova N.A., Arzamastsev A.A., Troitzsch K.G.* Development of a technology of designing intelligent information systems for the estimation of social objects // Institut für Wirtschafts- und Verwaltungsinformatik Fachbereich Informatik Universität Koblenz-Landau. № 1/2011. URL: http://www.uni-koblenz.de/~fb4reports/2011/2011_01_Arbeitsberichte.pdf. Загл. с экрана.
36. *Лифшиц В.Б., Булдакова Т.И., Калентьев С.В., Суетенков С.И.* Двухэтапное распознавание образов в нейросетевых диагностических системах // Информационные технологии в образовании, технике, медицине: международная научно-техническая конференция. 2000. С. 115-117.
37. *Арзамасцев А.А., Зенкова Н.А., Неудахин А.В.* Разработка экспертной системы с развивающимся интеллектуальным ядром на базе ИНС-моделей // Вестник Тамбовского университета. Серия Естественные и технические науки. Тамбов, 2010. Т. 15. Вып. 6. С. 1849-1857.
38. *Неудахин А.В., Арзамасцев А.А., Зенкова Н.А.* Построение нейросетевых экспертных систем с помощью автоматизированной технологии // Вопросы современной науки и практики. Университет им. В.И. Вернадского. 2009. № 4 (18). С. 179-183.
39. *Сернов С.П., Лифшиц В.Б., Субботина В.Г., Папицкая Н.Ю., Мартынова А.Г., Аредаков К.Г., Шульгин В.И.* Эпидемиология алкогольной болезни печени // Саратовский научно-медицинский журнал. 2009. Т. 5. № 4. С. 564-568.

БЛАГОДАРНОСТИ: Поддержано грантом Фонда содействия развития малых форм предприятий в научно-технической сфере. Проект «У.М.Н.И.К» № 14219: «Создание компьютерной программы прогноза течения алкогольной болезни печени» 2012 г.

Поступила в редакцию 5 апреля 2012 г.

Arzamastsev A.A., Lifshits V.B., Chichuk V.N. MATHEMATIC MODEL DEVELOPMENT OF ALCOHOLIC LIVER DISEASE COURSE PROGNOSIS

Mathematic model is a fast progressive scientific school. Alcoholic liver disease (ALD) is a topical problem of the present day. In the past no ways of modeling the ALD course were suggested. Some years ago the gap was successfully bridged using a cluster and discriminative analysis of considerable number of observations. In this work we have developed a mathematic model of the ALD course prognosis using another method, namely using the method of artificial neuron networks considering the small number of patients.

Key words: mathematic model; prognosis; alcoholic liver disease.