

СИНТЕЗ НЕЙРО-НЕЧЕТКИХ СЕТЕЙ С РАНЖИРОВАНИЕМ И СПЕЦИФИЧЕСКИМ КОДИРОВАНИЕМ ПРИЗНАКОВ ДЛЯ ДИАГНОСТИКИ И АВТОМАТИЧЕСКОЙ КЛАССИФИКАЦИИ ПО ПРЕЦЕДЕНТАМ

Решена задача автоматизации синтеза нейро-нечетких сетей для диагностирования и автоматической классификации по признакам. Предложен метод синтеза нейро-нечетких моделей по прецедентам, который определяет оценки взаимосвязи входных признаков, выделяет взаимосвязанные признаки, оценивает влияние признаков на выходной признак, а также выявляет взаимосвязанные термы, включает в модель наиболее важные признаки и термы, устраняет дуближ термов и признаков, использует различные варианты кодирования сигналов, а также переупорядочивает признаки, обеспечивая группировку признаков, формирует правила для прецедентов ранее не встречавшихся классов или существенно отличающихся от имеющихся правил своего класса, а уже имеющиеся правила корректирует на основе поступающих прецедентов с учетом числа ранее рассмотренных наблюдений. Предложенный метод позволяет существенно ускорить синтез нейро-нечетких моделей, обеспечивая приемлемую точность и более высокий уровень обобщения данных, снизить сложность и избыточность, а также повысить интерпретабельность нейромодели. Проведены эксперименты по решению практических задач диагностирования и автоматической классификации, подтвердившие работоспособность и применимость предложенного метода. Получена зависимость ошибки модели, синтезированной предложенным методом, от заданной разрядности признаков. Использование полученной зависимости позволяет на практике более рационально выбирать значение числа интервалов разбиения диапазонов значений признаков, обеспечивая приемлемую точность нейромодели.

Ключевые слова: выборка, нейро-нечеткая сеть, нечеткий вывод, обучение по прецедентам, диагностирование.

НОМЕНКЛАТУРА

ЭВМ – электронная вычислительная машина;
 δ_j – длина интервала, на которые разбивается диапазон значений j -го признака;
 μ_q – принадлежность к q -му классу;
 $\mu_{q,k}$ – принадлежность экземпляра к k -му эталону (правилу) q -го класса;
 $\mu_{q,k,i}$ – принадлежность экземпляра к k -му эталону (правилу) q -го класса по i -му признаку;
 $\mu_{q,k,i,l}$ – принадлежность экземпляра к k -му эталону (правилу) q -го класса по l -му интервалу i -го признака;
 $\mu_{i,l}(x^s)$ – функции принадлежности к l -му терму i -го признака;
 $\omega_{q,k}$ – нормированный вес k -го правила q -го класса;
 a, b, c, d – параметры функции принадлежности;
 C – набор правил-эталонов;
 $C^{(q,k)}$ – k -й эталон (правило) q -го класса;
 \bar{E} – ошибка модели;
 E – ошибка сети для задач с дискретным выходом;
 $E_{об.}$ – ошибка модели для обучающей выборки;
 $E_{тест.}$ – ошибка модели для тестовой выборки;
 f – критерий качества модели;
 $F()$ – структура нейромодели;
 $I^*(q,i,l)$ – значимость l -го интервала i -го признака относительно q -го класса;
 i, j – номера признаков;
 I_G – коэффициент обобщения нейромодели;

I_j – информативность j -го признака;
 $I_{i,l}$ – информативность l -го интервала i -го признака;
 I_w – число настраиваемых параметров нейромодели;
 k – номер правила (эталона);
 K – число классов;
 l – номер разряда (интервала);
 L – размерность кодирующего бинарного вектора – элемента сенсорной матрицы;
 N – число входных признаков;
 N' – число входных признаков, после отбора признаков;
 opt – условное обозначение оптимума;
 q – номер класса;
 Q^q – число правил (эталонов) q -го класса в наборе C ;
 $R(x^s, C^{(q,k)})$ – расстояние от экземпляра x^s до k -го эталона q -го класса $C^{(q,k)}$;
 $r_{i,j}$ – мера взаимосвязи входных признаков x_i и x_j ;
 $r_{i,y}$ – мера важности для каждого входного признака относительно выходного признака;
 S – объем выборки;
 $S_{об.}$ – число экземпляров в обучающей выборке;
 $S_{тест.}$ – число экземпляров в тестовой выборке;
 $t_{об.}$ – время, затраченное на обучение нейро-нечеткой модели;
 w – параметры нейромодели;
 $w^{(q,k)}$ – вес k -го правила q -го класса;
 x – набор входных векторов прецедентов;
 X – набор прецедентов;
 x_j – j -й входной признак;
 x^s – набор входных признаков s -го прецедента;
 x_j^s – значение j -го входного s -го прецедента;
 y – выходной признак;
 x_j^{\min}, x_j^{\max} – минимальное и максимальное значения признака x_j ;

$x_{j,l}^s$ – значение l -го разряда j -го признака s -го экземпляра в специфическом кодовом представлении;
 y^s – значение выходного признака для s -го прецедента.

ВВЕДЕНИЕ

Для обеспечения и поддержания высокого уровня качества, надежности и долговечности выпускаемых и эксплуатируемых изделий промышленного производства необходимо своевременно осуществлять их диагностирование [1].

Помимо методов физических измерений [1, 2], специфических для каждой конкретной задачи, ключевым элементом процесса диагностирования является диагностическая модель, построение которой требует либо полного знания физики диагностируемого объекта, либо привлечение человека-эксперта, либо использование методов машинного обучения [3].

Построение физических моделей [2], как правило, возможно только для хорошо изученных объектов диагностирования и специфически зависит от конкретной задачи. Такие методы непригодны для более широко встречающихся случаев, когда физика диагностируемого объекта является не полностью изученной, а сам объект подвергается воздействию внешних факторов [3].

Очевидными недостатками экспертных методов [4] является их чрезвычайная зависимость от субъективных знаний эксперта, его компетентности и ответственности, сложность извлечения и формализации знаний, а также сложность модификации и адаптации полученной модели принятия решений с учетом возможных изменений диагностируемого объекта, внешней среды.

Методы машинного обучения [5, 6], в отличие от методов построения физических моделей, не требуют знания физики диагностируемого объекта, а, в отличие от экспертных методов [4], не требуют привлечения человека-эксперта и способны автоматически извлекать знания из данных в процессе построения модели [3]. Поэтому данные методы целесообразно выбрать в качестве базиса для построения диагностических моделей.

Среди методов машинного обучения [5, 6] широкое распространение на практике в последнее время получили нейро-нечеткие сети [4, 7], которые способны обучаясь по прецедентам извлекать знания из данных, а также могут интегрировать в свою структуру экспертные знания и являются преобразуемыми в понятные для человека продукционные правила вида «если, то» [4].

Однако большинство известных методов построения нейро-нечетких моделей [4, 7] предполагают вовлечение пользователя в процесс создания сети и решение задач многомерной нелинейной оптимизации, характеризующихся неопределенностью выбора начальной точки поиска. Полученные диагностические модели в результате оказываются избыточными [3].

Объектом исследования данной работы является процесс построения диагностических моделей по прецедентам.

Предметом исследования данной работы являлись методы построения нейро-нечетких сетей для решения задач диагностирования.

Целью работы являлось создание метода синтеза диагностических моделей на основе нейро-нечетких сетей, свободного от указанных выше недостатков.

1 ПОСТАНОВКА ЗАДАЧИ

Пусть мы имеем исходную выборку $X = \langle x, y \rangle$ – набор S прецедентов о зависимости $y(x)$, $x = \{x^s\}$, $y = \{y^s\}$, $s = 1, 2, \dots, S$, характеризующихся набором N входных признаков $\{x_j\}$, $j = 1, 2, \dots, N$, и выходным признаком y . Каждый s -й прецедент представим как $\langle x^s, y^s \rangle$, $x^s = \{x_j^s\}$, $y^s \in \{1, 2, \dots, K\}$, $K > 1$.

Тогда задача синтеза нейромодели зависимости $y(x)$ будет заключаться в определении такой структуры $F()$ и значений параметров w нейромодели, при которых будет удовлетворен критерий качества модели $f(F)$, $w, \langle x, y \rangle \rightarrow opt$ [3]. Обычно критерий качества обучения нейромодели определяют как функцию ошибки модели:

$$\bar{E} = \frac{1}{2} \sum_{s=1}^S (y^s - F(w, x^s))^2 \rightarrow \min.$$

Для задач с дискретным выходом ошибку обученной сети можно характеризовать также формулой:

$$E = \frac{100\%}{S} \sum_{s=1}^S |y^s - F(w, x^s)| \rightarrow \min.$$

2 ЛИТЕРАТУРНЫЙ ОБЗОР

В настоящее время известно большое число различных архитектур нейро-нечетких сетей [4, 7]:

– нечеткие нейронные системы (fuzzy neural systems): в нейронных сетях применяются принципы нечеткой логики для ускорения процесса настройки или улучшения других параметров; нечеткая логика является лишь инструментом нейронных сетей и такая система не может быть интерпретирована в нечеткие правила, поскольку представляет собой «черный ящик»;

– конкурирующие нейро-нечеткие системы (concurrent neuro-fuzzy systems): нечеткая система и нейронная сеть работают над одной задачей, не влияя на параметры друг друга;

– параллельные нейро-нечеткие системы (cooperative neuro-fuzzy systems): настройка параметров выполняется с помощью нейронных сетей, после чего нечеткая система функционирует самостоятельно;

– интегрированные (гибридные) нейро-нечеткие системы (integrated neuro-fuzzy systems) – системы с тесным взаимодействием нечеткой логики и нейронных сетей. Данный тип нейро-нечетких сетей является наиболее подходящим для решения задач диагностирования.

Наиболее широко используемой на практике архитектурой интегрированных нейро-нечетких сетей являются сети Мамдани [8, 9], представляющие собой нейро-нечеткий классификатор. Однако методы их построения предполагают вовлечение пользователя в процесс создания сети (задание числа термов, видов их функций принадлежности, начальных значений параметров функций принадлежности термов, правил принятия решений), либо требуют решения задач кластер-анализа [10] набора прецедентов с последующей настройкой весов на основе градиентных методов оптимизации [11] с расчетом частных производных целевой функции на основе техники обратного распространения ошибки [4, 5].

Полученные диагностические модели в результате оказываются неоптимальными по структуре (зачастую избыточными), и, как следствие, оказываются менее точными и более сложными для восприятия и анализа человеком, что несколько осложняет их практическое применение [3].

Поэтому необходимо разработать метод, позволяющий получать диагностические модели, не требующие участия человека в процессе синтеза структуры и настройки параметров нейромодели, а также позволяющий снизить сложность получаемых моделей и повысить их интерпретируемость и обобщающие свойства.

3 МАТЕРИАЛЫ И МЕТОДЫ

Избыточность нейро-нечетких моделей, как правило, обусловлена тем, что набор используемых признаков является избыточным, а нечеткие термы формируются субъективно пользователем и могут излишне детализировать разбиение пространства признаков. Поэтому для снижения избыточности нейро-нечетких моделей предлагается в процессе синтеза сети определять оценки взаимосвязи входных признаков между собой, что позволит выделить взаимосвязанные признаки, оценивать влияние признаков на выходной признак (класс), а также выявлять взаимосвязанные термы, что позволит включать в модель наиболее важные признаки и термы, устранить дублирование термов и признаков путем оставления лишь части из них.

Для повышения вариативности (потенциала получения возможных решений) синтезируемых моделей предлагается использовать различные варианты кодирования сигналов, а также переупорядочивать признаки, обеспечивая тем самым группировку признаков по подобию с целью получения более легко воспринимаемых человеком диагностических моделей.

Для автоматизации формирования термов и выделения правил на основе заданного набора прецедентов предлагается реализовать процедуру адаптивного четкого кластер-анализа прецедентов, в процессе реализации которой формировать новые эталоны для прецедентов ранее не встречавшихся классов или существенно отличающихся от имеющихся эталонов своего класса, а уже имеющиеся эталоны корректировать на основе поступающих прецедентов с учетом числа ранее рассмотренных наблюдений.

На основе приведенных идей более формально изложим предлагаемый метод.

Этап инициализации. Задать обучающую выборку $\langle x, y \rangle$. Определить минимальное и максимальное значения признаков:

$$x_j^{\min} = \min_{s=1,2,\dots,S} \{x_j^s\}, x_j^{\max} = \max_{s=1,2,\dots,S} \{x_j^s\}.$$

Задать размерность кодирующего бинарного вектора – элемента сенсорной матрицы L . Очевидно, что должно быть $L > 1$.

Этап кодирования сигналов. Заменить значение каждого признака в выборке бинарным кодом. Для этого возможно использовать интервальное или колоночное кодирование.

Вначале на основе заданного L определим длину интервала каждого j -го признака:

$$\delta_j = \frac{x_j^{\max} - x_j^{\min}}{L}, j = 1, 2, \dots, N.$$

При интервальном кодировании диапазон значений каждого признака разбивается на заданное число интервалов равной длины, а значение признака, заменяется на бинарный вектор, элементы которого (разряды) соответствуют интервалам значений признака. В процессе кодирования определяется номер интервала, в который попало значение признака для конкретного экземпляра, после чего соответствующий элемент кодирующего вектора устанавливается равным единице, а остальные элементы – равными нулю:

$$x_{j,l}^s = \begin{cases} 0, l\delta_j < x_j^s - x_j^{\min}; \\ 1, (l-1)\delta_j \leq x_j^s - x_j^{\min} \leq l\delta_j; \\ 0, (l-1)\delta_j > x_j^s - x_j^{\min}, \end{cases}$$

$$j = 1, 2, \dots, N, l = 1, 2, \dots, L.$$

При колоночном кодировании диапазон значений каждого признака также разбивается на заданное число интервалов равной длины, значение признака, заменяется на бинарный вектор, элементы которого (разряды) соответствуют интервалам значений признака и определяется номер интервала, в который попало значение признака для конкретного экземпляра, однако все элементы кодирующего вектора, соответствующие данному интервалу и предыдущим ему интервалам устанавливаются равными единице, а остальные элементы, соответствующие интервалам больших значений признака – равными нулю:

$$x_{j,l}^s = \begin{cases} 1, l\delta_j \leq x_j^s - x_j^{\min}; \\ 0, l\delta_j > x_j^s - x_j^{\min}, \end{cases}$$

$$j = 1, 2, \dots, N, l = 1, 2, \dots, L.$$

Этап выделения правил. Каждый экземпляр выборки x^s в выбранном кодовом представлении $x^s = \{x_{j,l}^s\}$ представить в виде модифицированного продукционного правила $\{x_{i,l}^s\} \rightarrow y^s$.

Этап обобщения правил. Выполнить динамическое формирование правил-эталонов кластеров.

Для этого в начале задать пустой набор правил-эталонов $C = \emptyset$.

Затем просматривая экземпляры выборки ($s=1,2,\dots,S$) для текущего s -го экземпляра определить, имеется ли эталон его класса в наборе эталонов.

Если эталон для его класса отсутствует в C , то принять $q=y^s$, добавить эталон и занести в него s -й экземпляр: $C = C \cup C^{(q,1)}$, $C^{(q,1)} = x^s$: $C_{j,l}^{(q,1)} = \{x_{j,l}^s\}, j=1, 2, \dots, N; l=1, 2, \dots, L, w^{(q,1)}=1$.

В противном случае, если эталон класса для текущего s -го экземпляра отсутствует в C , то определить расстояние от экземпляра x^s до всех эталонов $R(x^s, C^{(q,k)})$, $q=1, 2, \dots, Q$; $k=1, 2, \dots, Q^q$:

$$R(x^s, C^{(q,k)}) = \sqrt{\sum_{j=1}^N \sum_{l=1}^L (x_{j,l}^s - C_{j,l}^{(q,k)})^2}.$$

Если расстояние от экземпляра x^s до ближайшего к нему эталона чужого класса меньше, чем расстояние от экземпляра до ближайшего эталона своего класса, т.е.

$$\min_{\substack{y^s \neq q, \\ k=1,2,\dots,Q^q}} \{R(x^s, C^{(q,k)})\} < \min_{\substack{y^s = q, \\ k=1,2,\dots,Q^q}} \{R(x^s, C^{(q,k)})\},$$

тогда сформировать новый эталон и занести в него экземпляр: $q=y^s$, $k=Q^q+1$, $C = C \cup C^{(q,k)}$, $C^{(q,k)} = \{x_{j,l}^s\}$, $w^{(q,k)}=1$; в противном случае – новый эталон получить как среднее между эталоном и экземпляром $C^{(q,k)} = (w^{(q,k)}C^{(q,k)} + x^s) / (w^{(q,k)} + 1)$:

$$C_{j,l}^{(q,k)} = \frac{w^{(q,k)}C_{j,l}^{(q,k)} + x_{j,l}^s}{w^{(q,k)} + 1},$$

$j=1, 2, \dots, N$; $l=1, 2, \dots, L$,
после чего принять: $w^{(q,k)} = w^{(q,k)} + 1$.

Этап ранжирования кодовых представлений сигналов. При построении и анализе диагностических и распознающих моделей желательно выявить и понимать взаимосвязь признаков, а также зависимость принимаемых решений от описательных признаков. Для этого предлагается определить ряд показателей.

Определим меру взаимосвязи входных признаков относительно интервалов (сенсоров) одного уровня, используя формулу:

$$r_{i,j} = r_{j,i} = \frac{\sum_{s=1}^S \sum_{l=1}^L \sum_{t=1}^L \{1 | x_{j,l}^s = x_{i,t}^s\}}{SL^2}, \quad i, j = 1, 2, \dots, N.$$

Признаки будут тем сильнее взаимосвязанными, чем больше экземпляров попадает в и интервалы (сенсоры) одного уровня.

Определим меру важности для каждого входного признака относительно выходного признака как

$$r_{i,y} = \frac{\sum_{s=1}^S \sum_{l=1}^L \sum_{t=1}^K \{1 | x_{j,l}^s = 1, y^s = t\}}{SLK}, \quad i = 1, 2, \dots, N.$$

Входной и выходной признаки будут тем сильнее взаимосвязанными, чем больше экземпляров попадает в и интервалы (сенсоры) одного класса.

После определения мер близости признаков ранжируем и перенумеруем признаки. В качестве первого признака целесообразно выбрать признак с наибольшим значением $r_{i,y}$. Затем в новый набор включим последовательно признаки, наиболее тесно связанные с предыдущими признаками, в порядке убывания значений $r_{i,y}$.

Данный этап можно пропустить для стационарных сигналов, распределенных во времени.

Этап оценки значимости интервалов и признаков. Признаки и их интервалы значений могут иметь различную ценность при принятии решений.

При этом для каждого класса значимость конкретного интервала признака определяется тем, сколько экземпляров, попавших в этот интервал будет принадлежать к данному классу и тем, сколько экземпляров, попавших в этот интервал не будет принадлежать к данному классу. Для определения значимости интервалов относительно классов предлагается использовать показатель, выражаемый формулой:

$$I^*(q, i, l) = 2I(q, i, l) - \sum_{p=1}^K I(p, i, l),$$

где $I(q, i, l) = \sum_{s=1}^S \sum_{l=1}^L \sum_{l=1}^L \{x_{i,l}^s | y^s = q\}$, $q=1, 2, \dots, K$; $i=1, 2, \dots, N$;
 $l=1, 2, \dots, L$.

На основе введенных показателей можно определить – показатель информативности l -го интервала i -го признака:

$$I_{i,l} = \frac{1}{K} \sum_{l=1}^K I^*(q, i, l), \quad i = 1, 2, \dots, N; \quad l = 1, 2, \dots, L;$$

– показатель информативности j -го признака:

$$I_i = I_{i,l} = \frac{1}{L} \sum_{l=1}^L I_{i,l} = \frac{1}{KL} \sum_{l=1}^L \sum_{q=1}^K I^*(q, i, l), \quad i = 1, 2, \dots, N.$$

Этап удаления малозначимых интервалов значений признаков. Все интервалы признаков, у которых значение $I_{i,l} = 0$, можно считать малозначимыми и удалить из дальнейшего рассмотрения. Среди оставшихся интервалов можно рассматривать как малозначимые те, для которых

$$I_{i,l} \leq \left[\frac{ES}{L} \right],$$

где E – заданное заранее значение максимально допустимой ошибки.

Этап удаления малозначимых признаков. Если среди признаков имеется такой, что все его интервалы признака малозначимыми, то данный признак следует исключить из рассмотрения.

Этап формирования термов признаков. На основе выделенных интервалов значений признаков возможно определить функции принадлежности к термам признаков на основе трапецевидной функции ($a \leq b \leq c \leq d$):

$$\mu_{i,l}(x^s) = \begin{cases} 0, & x \leq a; \\ \frac{x-a}{b-a}, & a \leq x \leq b; \\ 1, & b \leq x \leq c; \\ \frac{d-x}{d-c}, & c \leq x \leq d; \\ 0, & d \leq x. \end{cases}$$

При інтервальному кодированні параметри функції приналежності к l -му терму i -го признака будуть ініціалізуватися по формулам:

$$a = (l-1)\delta + x_i^{\min}, b = (l-1)\delta + x_i^{\min}, c = l\delta + x_i^{\min}, \\ d = l\delta + x_i^{\min}.$$

При колоночному кодированні параметри функції приналежності к l -му терму i -го признака будуть ініціалізуватися по формулам:

$$a = x_i^{\min}, b = x_i^{\min}, c = l\delta + x_i^{\min}, d = l\delta + x_i^{\min}.$$

Етап формирования системы нечеткого вывода. На данном этапе задаются принципы преобразования значений функций принадлежности к термам в принадлежности к кластерам и классам, а также способ дефаззификации результата.

Принадлежности к сформированным эталонам определим по формулам:

$$\mu^{q,k} = \max_{i=1,2,\dots,N} \{\mu_{q,k,i}\} \text{ либо } \mu^{q,k} = \max_{i=1,2,\dots,N} \left\{ \frac{I_i \mu_{q,k,i}}{\sum_{j=1}^N I_j} \right\},$$

где

$$\mu_{q,k,i} = \max_{l=1,2,\dots,L} \{\mu_{q,k,i,l}\} \text{ либо}$$

$$\mu_{q,k,i} = \max_{l=1,2,\dots,L} \left\{ \frac{I_{i,l} \mu_{q,k,i,l}}{\sum_{p=1}^L I_{i,p}} \right\},$$

$$\mu_{q,k,i,l} = \mu_{i,l}(x^s) C_{i,l}^{(q,k)},$$

$$q=1, 2, \dots, K, k=1, 2, \dots, Q^q, i=1, 2, \dots, N; l=1, 2, \dots, L.$$

Принадлежности к классам определим формулами:

$$\mu_q = \max_{k=1,2,\dots,Q^q} \{\omega_{q,k} \mu^{q,k}\},$$

$$\text{где } \omega_{q,k} = \frac{w^{(q,k)}}{\sum_{p=1}^{Q^q} w^{(q,p)}}.$$

Номер класса, соответственно, будет определяться по формуле:

$$y^s = \arg \max_{q=1,2,\dots,K} \{\mu_q\} \text{ или } y^s = \left[\frac{\sum_{q=1}^K q \mu_q}{\sum_{q=1}^K \mu_q} \right].$$

Этап синтеза нейро-нечеткой сети. На основе выделенных термов и определенных функций системы нечеткого вывода можно сформировать структуру нейро-нечеткой сети в соответствии со схемой, приведенной на рис. 1.

На входной (первый) слой сети поступают входные сигналы – значения признаков распознаваемого экземпляра, которые далее поступают на второй слой, выполняющий фаззификацию. Выходы нейронов второго слоя представляют собой значения функций принадлежности распознаваемого экземпляра к термам признаков. Нейроэлементы третьего слоя сети комбинируют принадлежности к термам в принадлежности к эталонам кластеров, на основе которых нейроны четвертого слоя определяют принадлежности к классам. Единственный нейрон пятого (выходного) слоя осуществляет дефаззификацию и выдает на выходе расчетное значение номера класса распознаваемого экземпляра.

Этап настройки весов. На основе обучающей выборки оценить ошибку для построенной сети. Если ошибка является приемлемой, то завершить синтез сети, в противном случае выполнить процедуру обучения посредством решения задачи многомерной оптимизации параметров нечетких термов второго слоя, используя метод обратного распространения ошибки [4, 11] или эволюционные алгоритмы [3, 5].

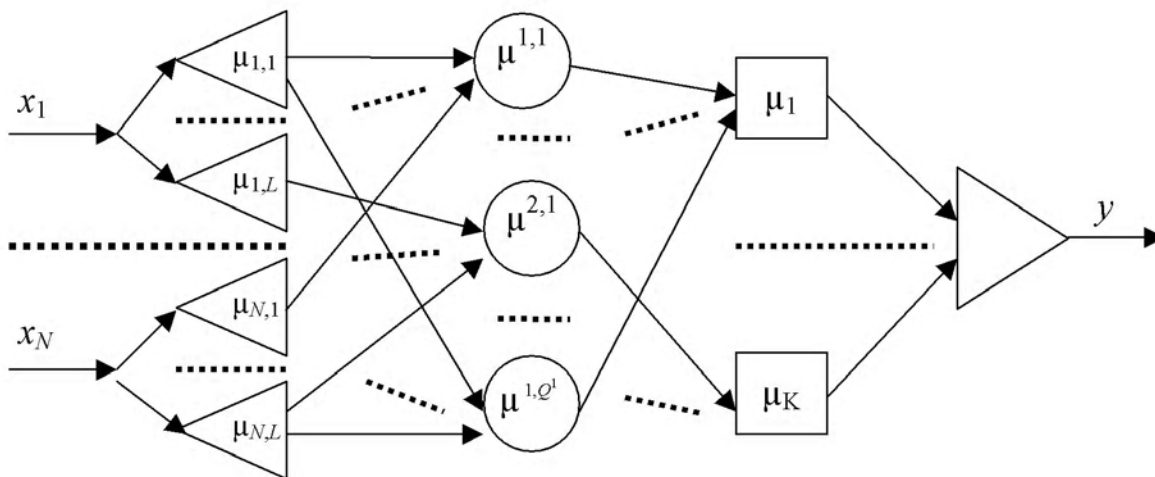


Рисунок 1 – Схема нейро-нечеткой сети

Если в результате обучения достигнут требуемый уровень точности (приемлемая ошибка), то завершить синтез сети, в противном случае – увеличить число интервалов L и перейти к этапу кодирования. Поскольку число интервалов L не должно превышать числа экземпляров в выборке S , то при получении на данном этапе $L \geq S$ следует прекратить дальнейший поиск и вернуть полученную модель с наилучшей точностью и наименьшим L .

4 ЭКСПЕРИМЕНТЫ

Для исследования практической применимости предложенного метода он был программно реализован как дополнение к Автоматизированной системе синтеза нейросетевых и нейро-нечетких моделей для неразрушающей диагностики и классификации образов по признакам [12].

Разработанное программное обеспечение использовалось для решения задач неразрушающего диагностирования и автоматической классификации, характеристики которых представлены в табл. 1.

Для каждой задачи на основе кластер анализа человеком выделялись нечеткие термы и правила, по которым определялась структура нейро-нечетких моделей, после чего на основе обратного распространения ошибки выполнялась настройка параметров моделей. Также для каждой задачи на основе выборки данных в автоматическом режиме на основе предложенного метода синтезировались нейро-нечеткие модели.

5 РЕЗУЛЬТАТЫ

Результаты проведенных вычислительных экспериментов представлены в табл. 2.

Здесь коэффициент обобщения определялся по формуле:

$$I_G = \frac{NS}{I_w}, I_w \geq 1.$$

Чем больше значение коэффициента обобщения, тем выше уровень обобщения обучающей выборки данных нейро-нечеткой моделью.

С целью изучения зависимости ошибки модели E от заданной разрядности признаков L для различных задач (выборок данных различной размерности и природы [3]) получен график, приведенный на рис. 2.

На рис. 3 представлен график усредненной зависимости ошибки модели E от доли отобранных признаков в первоначальном наборе признаков N'/N для различных задач (выборок данных различной размерности и природы [3]).

На рис. 4 представлен график усредненной зависимости коэффициента обобщения модели I_G от доли отобранных признаков в первоначальном наборе признаков N'/N для различных задач (выборок данных различной размерности и природы [3]).

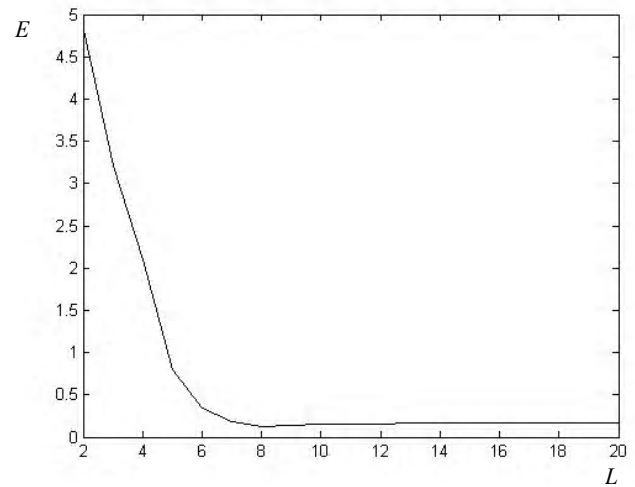


Рисунок 2 – График усредненной зависимости E от L

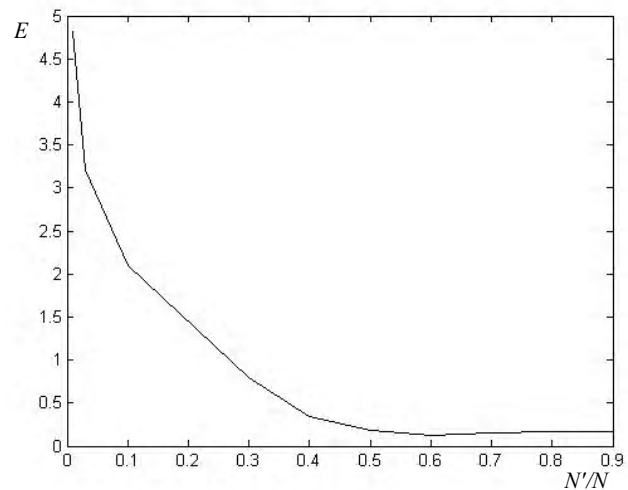


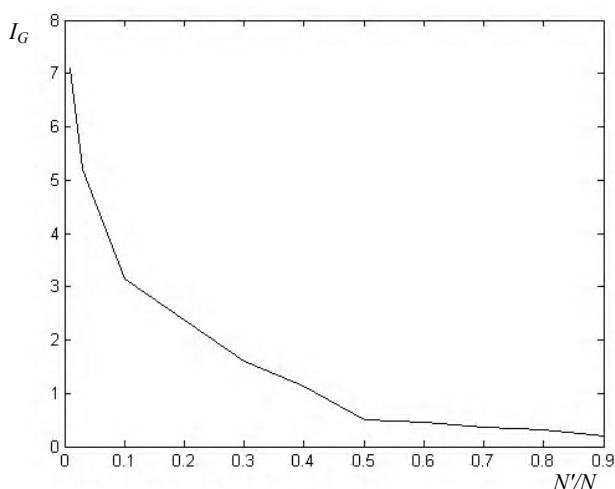
Рисунок 3 – График усредненной зависимости E от N'/N

Таблица 1 – Характеристики практических задач диагностирования и автоматической классификации

Название задачи	N	$S_{об.}$	$S_{тест.}$	K
Диагностирование лопаток газотурбинных авиадвигателей по спектрам свободных затухающих колебаний после ударного возбуждения [3]	100	16	16	2
Автоматическая классификация растительных объектов по коэффициентам спектральной яркости [13]	55	120	128	2
Диагностирование хронического обструктивного бронхита [14]	28	102	103	2

Таблица 2 – Результаты вычислительных экспериментов

Название задачи	Сеть Мамдани, заданная человеком и обученная методом обратного распространения ошибки				Сеть Мамдани, синтезированная на основе предложенного метода			
	$E_{об.}, \%$	$E_{тест.}, \%$	$t_{об.}, с$	I_G	E	$E_{тест.}$	$t_{об.}, с$	I_G
Диагностирование лопаток газотурбинных авиадвигателей [3]	0	0	13,92	0,8	0	0	12,75	1,33
Автоматическая классификация растительных объектов [13]	1,7	2,3	23,14	4,29	0	0,8	21,89	5
Диагностирование хронического обструктивного бронхита [14]	4,9	2,9	18,85	2,83	0	0,9	17,14	3,19

Рисунок 4 – График усредненной зависимости I_G от N'/N

6 ОБСУЖДЕНИЕ

Как видно из табл. 2, предложенный метод позволяет существенно ускорить синтез нейро-нечетких моделей, обеспечивая приемлемую точность и более высокий уровень обобщения данных моделью. Это позволяет рекомендовать предложенный метод для использования на практике.

Полученная зависимость ошибки модели E от заданной разрядности признаков L для выборок данных различной размерности и природы, график которой приведен на рис. 2, показывает, что при начальном значении $L=2$ ошибка также имеет наибольшее значение, которое постепенно снижается до некоторого значения, после чего начинается постепенный рост ошибки. Изначально высокое значение ошибки объясняется недостаточной разрядностью кодирования вследствие большой потери информации при разбиении на интервалы значений признаков из-за малого числа интервалов. Постепенное убывание ошибки по мере роста значения L объясняется увеличением объема информации, доступного для отображения в модели. Дальнейший рост ошибки объясняется ухудшением обобщающих свойств модели вследствие избыточности ее структуры и параметров, а также излишней детализации разбиения признакового пространства.

Представленный на рис. 3 график показывает, что при небольшом числе используемых признаков относительно размера исходного набора признаков ошибка построенной модели является большей, чем при использовании большего числа признаков. Однако при приближении числа используемых в модели признаков к числу признаков исходного набора убывание ошибки становится незначительным. Таким образом, отбор признаков должен быть направлен на поиск оптимального соотношения N'/N , обеспечивающего приемлемый уровень ошибки.

График, изображенный на рис. 4, показывает, что при большом числе используемых признаков относительно размера исходного набора признаков коэффициент обобщения построенной модели является большим, чем при использовании большего числа признаков. Таким образом, отбор признаков должен быть направлен также на поиск оптимального соотношения N'/N , обеспечивающего не только приемлемый уровень ошибки, но и высокий уровень обобщения модели.

ЗАКЛЮЧЕНИЕ

В работе решена актуальная задача автоматизации синтеза нейро-нечетких сетей для диагностирования и автоматической классификации по признакам.

Научная новизна работы состоит в том, что предложен метод синтеза нейро-нечетких моделей по прецедентам, который в

процессе синтеза сети определяет оценки взаимосвязи входных признаков между собой, выделяет взаимосвязанные признаки, оценивает влияние признаков на выходной признак, а также выявляет взаимосвязанные термы, включает в модель наиболее важные признаки и термы, устраняет дуближ термов и признаков путем оставления лишь части из них, использует различные варианты кодирования сигналов, а также переупорядочивает признаки, обеспечивая группировку признаков, формирует правила для прецедентов ранее не встречавшихся классов или существенно отличающихся от имеющихся правил своего класса, а уже имеющиеся правила корректирует на основе поступающих прецедентов с учетом числа ранее рассмотренных наблюдений.

Предложенный метод позволяет существенно ускорить синтез нейро-нечетких моделей, обеспечивая приемлемую точность и более высокий уровень обобщения данных, снизить сложность и избыточность, а также повысить интерпретируемость нейромодели.

Практическая значимость результатов работы состоит в том, что проведены эксперименты по решению практических задач диагностирования и автоматической классификации, подтвердившие работоспособность и применимость предложенного метода. Получена зависимость ошибки модели, синтезированной предложенным методом, от заданной разрядности признаков. Использование полученной зависимости позволяет на практике более рационально выбирать значение числа интервалов разбиения диапазонов значений признаков, обеспечивая приемлемую точность нейромодели. Получены зависимости ошибки синтезированной модели и коэффициента обобщения от доли используемых в модели признаков в исходном наборе признаков.

Перспективы дальнейших исследований заключаются в том, чтобы изучить зависимость оптимального значения числа интервалов разбиения диапазонов значений признаков от информативности набора признаков и размера набора признаков.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке международного проекта «Centers of Excellence for young REsearchers» программы «Темпус» Европейской Комиссии (№ 544137-TEMPUS-1-2013-1-SK-TEMPUS-JPHES) в рамках госбюджетных научно-исследовательских тем Запорожского национального технического университета «Интеллектуальные информационные технологии диагностирования и автоматической классификации» и «Интеллектуальные методы диагностирования систем управления удаленными техническими объектами».

СПИСОК ЛИТЕРАТУРЫ

1. Биргер И. А. Техническая диагностика / И. А. Биргер. – М. : Машиностроение, 1978. – 240 с.
2. Ding S. X. Model-based fault diagnosis techniques: design schemes, algorithms, and tools / S. X. Ding. – Berlin: Springer, 2008. – 473 p.
3. Интеллектуальные информационные технологии проектирования автоматизированных систем диагностирования и распознавания образов : монография / [С. А. Субботин, Ан. А. Олейник, Е. А. Гофман, С. А. Зайцев, Ал. А. Олейник] ; под ред. С. А. Субботина. – Харьков : Компания СМІТ, 2012. – 318 с.
4. Субботин С.О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : навч. посібник / С. О. Субботин. – Запоріжжя : ЗНТУ, 2008. – 341 с.
5. Computational intelligence: a methodological introduction / [R. Kruse, C. Borgelt, F.Klawonn et. al.]. – London: Springer-Verlag, 2013. – 488 p. DOI: 10.1007/978-1-4471-5013-8_1
6. Bishop C. M. Pattern recognition and machine learning / C. M. Bishop. – New York : Springer, 2011. – 740 p.
7. Buckleya J. J. Fuzzy neural networks: a survey / J. J. Buckleya, Y. Hayashi // Fuzzy sets and systems. – 1994. –Vol. 66, Issue 1. – P. 1–13.

8. Mamdani E. H. An experiment in linguistic synthesis with fuzzy logic controller / E. H. Mamdani, S. Assilian // *International journal of man-machine studies*. – 1975. – Vol. 7, № 1. – P. 1–13.
9. Chai Y. Mamdani Model based Adaptive Neural Fuzzy Inference System and its Application / Y. Chai, L. Jia, Z. Zhang // *International Journal of Computer, Electrical, Automation, Control and Information Engineering* Vol:3, No:3, 2009. –P. 663–670.
10. Леоненков А. В. Нечеткое моделирование в среде MATLAB и fuzzyTECH / А. В. Леоненков. – СПб.: БХВ-Петербург, 2003. – 736 с.
11. Химмельблау Д. Прикладное нелинейное программирование / Д. Химмельблау. – М.: Мир, 1974. – 534 с.
12. Комп'ютерна програма «Автоматизована система синтезу нейромережних та нейро-нечітких моделей для неруйнівної

Субботін С. О.

Д-р техн. наук, професор, професор кафедри програмних засобів Запорізького національного технічного університету, Запоріжжя, Україна

СИНТЕЗ НЕЙРО-НЕЧІТКИХ МЕРЕЖ З РАНЖИРУВАННЯМ І СПЕЦИФІЧНИМ КОДУВАННЯМ ОЗНАК ДЛЯ ДІАГНОСТИКИ Й АВТОМАТИЧНОЇ КЛАСИФІКАЦІЇ ЗА ПРЕЦЕДЕНТАМИ

Вирішено задачу автоматизації синтезу нейро-нечітких мереж для діагностування й автоматичної класифікації за ознаками. Запропоновано метод синтезу нейро-нечітких моделей за прецедентами, що визначає оцінки взаємозв'язку вхідних ознак, виділяє взаємозалежні ознаки, оцінює вплив ознак на вихідну ознаку, а також виявляє взаємозалежні терми, включає в модель найбільш важливі ознаки і терми, усуває дубляж термів і ознак, використовує різні варіанти кодування сигналів, а також переупорядковує ознаки, забезпечуючи групування ознак, формує правила для прецедентів класів, що раніше не зустрічалися, або таких, що істотно відрізняються від наявних правил свого класу, а уже наявні правила коректує на основі прецедентів, що надходять, з урахуванням числа раніше розглянутих спостережень. Запропонований метод дозволяє істотно прискорити синтез нейро-нечітких моделей, забезпечуючи прийнятну точність і більш високий рівень узагальнення даних, знизити складність і надмірність, а також підвищити інтерпретабельність нейромоделі. Проведено експерименти з вирішення практичних задач діагностування й автоматичної класифікації, що підтвердили працездатність і застосовність запропонованого методу. Отримано залежність помилки моделі, синтезованої запропонованим методом, від заданої розрядності ознак. Використання отриманої залежності дозволяє на практиці більш раціонально вибирати значення кількості інтервалів розбиття діапазонів значень ознак, забезпечуючи прийнятну точність нейромоделі.

Ключові слова: вибірка, нейро-нечітка мережа, нечітке виведення, навчання за прецедентами, діагностування.

Subbotin S. A.

Dr.Sc., Professor, Professor of department of software tools, Zaporizhzhya National Technical University, Zaporizhzhya, Ukraine

THE NEURO-FUZZY NETWORK SYNTHESIS WITH THE RANKING AND SPECIFIC ENCODING OF FEATURES FOR THE DIAGNOSIS AND AUTOMATIC CLASSIFICATION ON PRECEDENTS

The problem of automation synthesis of neuro-fuzzy networks for diagnostics and automatic classification on features is solved. The method of neuro-fuzzy model synthesis on precedents is proposed.

It evaluates the relationship of input features, extracts related features, evaluates the impact of features on the output feature and identify related terms, includes the most important features and terms in the model, eliminates dubbing of the terms and features, uses different types of signal encoding, and also arranges features providing their grouping, forms a rules for precedents of previously non-experienced classes or significantly different from the existing rules of their class, adjust the existing rules based on incoming precedents, given the number previously considered observations. The proposed method allows to significantly accelerate the synthesis of neuro-fuzzy models, providing acceptable accuracy and a higher level of generalization of data, reduce complexity and redundancy, as well as increase of interpretability of neural model. The experiments on solution of practical problems of diagnosis and automatic classification are conducted. They confirmed the efficiency and applicability of the proposed method. The dependence of error of the model synthesized by the proposed method from the specified feature bitness is obtained. Using the obtained dependence allows to more rational choose the value of the number of divisions of the range of feature values in practice providing a reasonably accurate neural model.

Keywords: sample, neuro-fuzzy network, fuzzy inference, training on precedents.

REFERENCES

1. Birger I. A. *Tekhnicheskaya diagnostika*. Moscow. Mashinostroyeniye, 1978, 240 p.
2. Ding S. X. *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Berlin, Springer, 2008, 473 p.
3. Subbotin S. A., Oleynik An. A., Gofman Ye. A., Zaytsev S. A., Oleynik Al. A. ; pod red. S. A. Subbotina. *Intellektual'nyye informatsionnyye tekhnologii proyektirovaniya avtomatizirovannykh sistem diagnostirovaniya i raspoznavaniya obrazov* : monografiya. Khar'kov, Kompaniya SMIT, 2012, 318 p.
4. Subbotin S. O. *Podannya u obrobka znan' u sistemakh shuchnoho intelektu ta pidtrymky pryynyattya rishen'* : navch. posibnyk. Zaporizhzhya, ZNTU, 2008, 341 p.
5. Kruse R., Borgelt C., Klawonn F. et. al. *Computational intelligence: a methodological introduction*. London: Springer-Verlag, 2013, 488 p. DOI: 10.1007/978-1-4471-5013-8_1
6. Bishop C. M. *Pattern recognition and machine learning*. New York, Springer, 2011, 740 p.
7. Buckleya J. J., Hayashi Y. *Fuzzy neural networks: a survey*, *Fuzzy sets and systems*, 1994, Vol. 66, Issue 1, pp. 1–13.
8. Mamdani E. H. Assilian S. An experiment in linguistic synthesis with fuzzy logic controller, *International journal of man-machine studies*, 1975, Vol. 7, No. 1, pp. 1–13.
9. Chai Y., Jia L., Zhang Z. Mamdani Model based Adaptive Neural Fuzzy Inference System and its Application, *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, Vol:3, No:3, 2009, pp. 663–670.
10. Leonenkov A. V. *Nechotkoye modelirovaniye v srede MATLAB i fuzzyTECH*. Sankt-Peterburg, BKHV-Peterburg, 2003, 736 p.
11. Khimmel'blau D. *Prikladnoye nelineynoye programmirovaniye*. Moscow, Mir, 1974, 534 p.
12. Subbotin S. O. *Komp'yuterna prohrama «Avtomatyrovana sistema syntezu neyromerezhevykh ta neyro-nechitkykh modeley dlya neruynivnoyi diahnostryky ta klasyfikatsiyi obraziv za oznakamy»* : cvidotstvo pro reyestratsiyu avtors'koho prava na tvir № 35431. Derzh. departament intellektual'noyi vlasnosti. № 34011 ; zayavl. 21.04.10 ; zareyestr. 21.10.10.
13. Shama Ye. O., Subbotin S. O., Morshevavka S. V. *Pobudova klasyfikatora roslynnykh ob'yektiv za dopomohoyu neyronnykh merezh*, *Radioelektronika, informatyka, upravlinnya*, 2013, No. 1, pp. 55–61.
14. Kolisnyk N. V., Subotyn S. O. *Modelyuvannya imunopatohezezu khronichnoho obstruktyvnoho bronkhitu za dopomohoyu neyromerezh*. Suchasni problemy biolohiyi, ekolohiyi ta khimiyi : II Mizhnarodna konferentsiya, Zaporizhzhya, 1–3 zhovtnya 2009 r. : zbirka materialiv. Zaporizhzhya, ZNU, pp. 124–125.

Статья поступила в редакцию 17.01.2016.

После доработки 10.02.2016.