

ИНФОРМАТИКА,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И УПРАВЛЕНИЕ

INFORMATICS,
COMPPUTER ENGINEERING
AND CONTROL

УДК 004.852

DOI: 10.17212/1814-1196-2018-3-73-86

Нейро-нечеткая классификация объектов и их состояний^{*}

А.М. КОРИКОВ^{1,2,а}, А.Т. НГУЕН^{2,б}

¹ 634050, РФ, г. Томск, пр. Ленина, 40, Томский государственный университет систем управления и радиоэлектроники

² 634050, РФ, г. Томск, пр. Ленина, 30, Национальный исследовательский Томский политехнический университет

^а korikov@asu.tusur.ru ^б nguyenanhtu@tpu.ru

В работе представлена разработка и исследование нейронной нечеткой сети (ННС) для решения задач классификации биологических объектов и их состояний. Рассмотрен общий подход к задачам автоматической классификация объектов. Постановка задач классификации объектов представлена на типовом наборе данных ирисы Фишера, а оценка эмоционального состояния объектов – на наборе данных шкалы баланса. ННС является комбинацией систем нечеткого вывода и нейронной сети. Структура ННС представляет собой многослойную одно-направленную сеть, состоящую из следующих слоев: входной слой, слой нечетких функций активации (ФА), слой фаззификации, слой дефаззификации, слой нормализации и выходной слой. В слое нечетких ФА используются нечеткие ФА, разработанные и исследованные с участием авторов. Представлены четыре вида нечетких ФА, поэтому разрабатываются и исследуются четыре вида ННС. Обучение ННС – определение antecedentных параметров сети – осуществляется с помощью метода кластеризации К-средних. В обучении ННС используется алгоритм масштабированного сопряженного градиента (SCG), который позволяет снизить трудоемкость поиска линий на каждую итерацию обучения, что повышает его быстродействие. Тестирование ННС выполнено на наборах данных ирисы Фишера и шкалы баланса. Эти наборы данных являются классическими и часто используются на практике для иллюстрации работы различных статистических алгоритмов классификации. Эксперименты с разными видами нечетких ФА подтверждают эффективность применения ННС для решения задач классификации. Выполнен сравнительный анализ ННС предлагаемых четырех видов и известной ННС с гауссовой ФА по эффективности и точности классификации.

Ключевая слова: нейронная нечеткая сеть, классификация, нечеткие функции активации, алгоритм масштабированного сопряженного градиента, ирисы Фишера, шкалы баланса, эффективность классификации, точность классификации

^{*} Статья получена 18 апреля 2018 г.

ВВЕДЕНИЕ

Старт исследованиям по теории нейронных сетей (НС) и внедрению нейросетевых технологий (НСТ) во многие направления науки и техники был дан в 1943 г. работой У. МакКаллока и У. Питса [1]. Теория НС и НСТ развивается в рамках обширнейшей области науки и техники – искусственного интеллекта (ИИ) [2]. В естественных и технических науках задачи делятся на формализуемые и неформализуемые. Постановка любой задачи заключается в том, чтобы перевести ее словесное, вербальное описание в формальное. В случае относительно простых задач такой переход осуществляется в сознании (естественном интеллекте (ЕИ)) человека. Если полученная формальная модель (математическая зависимость между величинами в виде формулы, уравнения, системы уравнений) опирается на фундаментальный закон или подтверждается экспериментом, то этим доказывается ее адекватность отображаемой ситуации, и модель рекомендуется для решения задач соответствующего класса [3]. Неформализуемые задачи (НФ-задачи) невозможно представить, используя только ЕИ человека, в виде каких-то формальных моделей. С развитием науки и техники количество таких НФ-задач постоянно растет. Для их решения развиваются специальные приемы и методы, обобщенные в теории систем и системном анализе [3], создаются экспертные системы – интеллектуальные системы (ИС), основанные на знаниях [4]. Роль методов ИИ, теории НС и НСТ в реализации методик системного анализа и разработке ИС трудно переоценить. Решение НФ-задач осложняется различными видами НЕ-факторов знаний [4]: неопределенность, нечеткость, неточность, недоопределенность и неполнота знаний эксперта о свойствах проблемной области.

Спектр задач, решаемых на основе НСТ, постоянно расширяется. Некоторое время считалось, что НСТ эффективны при решении трудноформализуемых задач и НФ-задач, для которых необходимо включение в алгоритм их решения процесса обучения на результатах реального эксперимента [1]. В основном это были задачи распознавания образов. В последние годы к ним добавились задачи обработки сигналов и изображений, распараллеливание которых повышает эффективность (например, скорость их решения). Эти задачи не требуют обучения на результатах реального эксперимента, но они хорошо представимы в нейросетевом логическом базисе [1].

В настоящей работе решаются задачи автоматической классификации объектов и их состояний. Общий подход к задачам автоматической классификации объектов изложен в [5]. Согласно этому подходу эмпирические данные представляются в виде матрицы данных Z , строки которой соответствуют различным наблюдаемым объектам, а столбцы – параметрам, описывающим состояние каждого объекта. Если параметры имеют различный физический смысл, то осуществляется преобразование матрицы данных Z в стандартизованную матрицу данных X . Формулы преобразования Z в X приведены в [5]. Матрице данных ставится в соответствие корреляционная матрица R , элементами которой являются выборочные коэффициенты корреляции r_{ij} , пропорциональные скалярному произведению двух вектор-столбцов матрицы данных – скалярное произведение i -го вектор-столбца и j -го вектор-столбца. Коэффициент корреляции является удобным показате-

лем «близости» или «связи» между параметрами. На этой основе разработаны многие структурные методы обработки эмпирических данных, в том числе методы автоматической классификации объектов, образов и изображений [5]. Задача классификации объектов трактуется как геометрическая задача о выделении в пространстве классификационных признаков «плотных» скоплений точек. Для ее решения необходимо задание правила отнесения любого объекта к тому или иному классу. Подобные задачи сформулированы и решены в многочисленных отечественных и зарубежных публикациях. Однако поиск эффективных решений задач классификации объектов и их состояний продолжается. В настоящей работе для решения задач классификации объектов и их состояний разрабатывается и исследуется нечеткая НС, базирующаяся на нечетких функциях активации (ФА), предложенных авторами в [6, 7], и алгоритме масштабированного сопряженного градиента (SCG) [8].

1. ПОСТАНОВКА ЗАДАЧИ

Нейронная сеть представляет собой сложные отношения между входами и выходами [1, 2]. На основе объединения результатов теории нейронных сетей и теории нечетких множеств возможно создание оригинальных интеллектуальных систем [10]. Нечеткие системы, как правило, не учатся и не приспосабливаются [11], в то время как нейронные сети обладают способностью к самообучению и самоорганизации. По этим причинам вполне логично создание нейро-нечетких систем, которые представляют собой комбинацию систем нечеткого вывода и нейронной сети. Нейро-нечеткая система представляет собой нечеткую нейронную сеть, которая включает в себя систему нечетких выводов, преодолевает некоторые ограничения нейронных сетей и расширяет пределы нечетких систем. В известных нечетких нейронных системах – нейро-нечетких классификаторах (ННК) – каждый узел сети описывается параметризованной колоколообразной функцией активации (ФА) или гауссовой ФА. В наших работах [6, 7] предлагаются новые модели нечетких НС (ННС), использующих импульсные нечеткие ФА нейронов ННС. Поставим задачу разработки ННК на основе этих нейронов, обучения ННК алгоритма масштабированного сопряженного градиента (SCG) и исследования возможности ННК для решения задач классификации объектов и их состояний. Для тестирования ННК выберем классические наборы данных: ирисы Фишера и шкалы баланса. Ирисы Фишера – это набор данных для задачи классификации, на примере которого Рональд Фишер в 1936 году продемонстрировал эффективность разработанного им метода дискриминантного анализа [11]. Этот набор данных стал уже классическим и часто используется для иллюстрации эффективности работы различных статистических алгоритмов. Шкалы баланса также являются классическим набором данных для задачи классификации. Этот набор данных предложен Зиглером [12] и описывает процесс взвешивания с использованием шкалы баланса. Исследование проведено с помощью средств компьютерного моделирования MATLAB [13].

2. НЕЙРО-НЕЧЕТКИЙ КЛАССИФИКАТОР

Задача нечеткой классификации – это задача разбиения пространства классификационных признаков на нечеткие классы. Можно описать пространство классификационных признаков с нечеткими областями и управлять каждой областью с помощью нечетких правил [14]. Типичное нечеткое правило классификации P_i , которое демонстрирует связь между пространством входных данных и классами, выглядит следующим образом:

$$P_i: \text{если } x_{s1} - A_{i1} \text{ и } \dots x_{sj} - A_{ij} \dots \text{ и } x_{sn} - A_{in}, \text{ то класс} - C_k,$$

где x_{sj} представляет j -й входной элемент s -й выборки; A_{ij} обозначает нечеткое множество j -го входного элемента в i -м правиле; C_k представляет собой метку k -го класса. A_{ij} определяется соответствующей функцией принадлежности.

Для достижения этой цели нечеткий параметр функции принадлежности [9] и нечеткий кластер должны быть оптимизированы. Когда правила нечеткой классификации строятся как сеть, то их параметры могут быть адаптированы к нейронной сети. В результате системы с нечеткой классификацией и нейронные сети могут быть объединены с сохранением их положительных свойств. Комбинированная система называется нейро-нечетким классификатором, который представляет собой адаптивную сеть с несколькими входами и несколькими выходами [16].

Нейро-нечеткий классификатор представляет собой многослойную одностороннюю сеть, состоящую из следующих слоев: входной слой, слой нечетких ФА, слой фазификации, слой дефазификации, слой нормализации и выходной слой. Классификатор имеет несколько входов и несколько выходов. На рис. 1 изображен ННК с двумя классификационными признаками и тремя классами. Каждый вход определяется тремя лингвистическими переменными. Таким образом, существует девять нечетких правил.

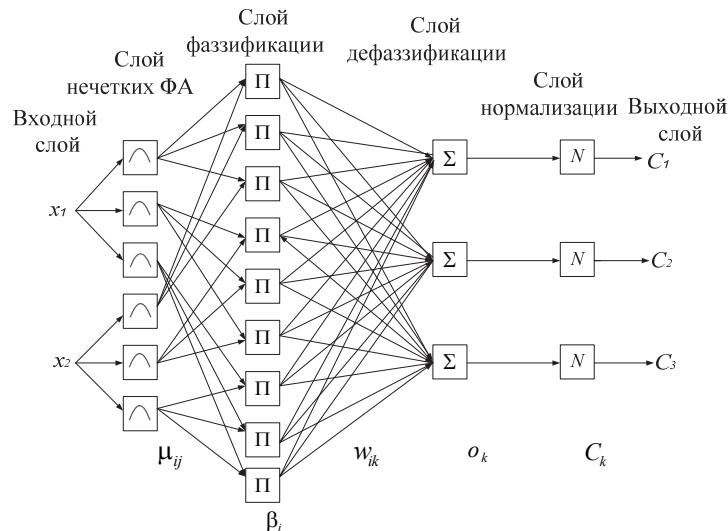


Рис. 1. Структура нейро-нечеткого классификатора

Fig. 1. The structure of a neuro-fuzzy classifier

Слой нечетких ФА. Функция активации для каждого входа идентифицируется в этом слое. В данной работе используется четыре вида нечетких ФА [6, 7]. На рис. 2 показан внешний вид этих ФА.

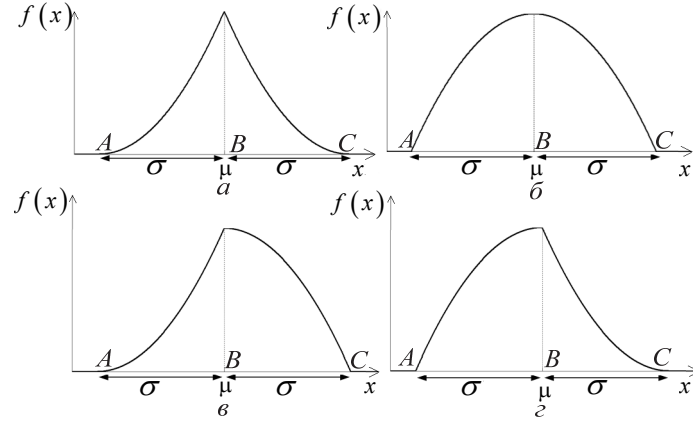


Рис. 2. Внешний вид ФА

Fig. 2. The appearance of AF

Набор точек $\langle A, B, C \rangle$ является характеристическими точками ФА. Задаем $AB = BC = \sigma$ и $B = \mu$. Пусть $a = (\mu - \sigma)$, $k_1 = 0$, $k_2 = 1$, если $\mu - \sigma \leq x \leq \mu$ и $a = (\mu + \sigma)$, $k_1 = 1$, $k_2 = 0$, если $\mu \leq x \leq (\mu + \sigma)$, тогда получим 4 ФА соответственно:

$$\begin{aligned}
 a) \quad & \begin{cases} f_{ij}(x_{sj}) = \frac{1}{\sigma_{ij}^2} x_{sj}^2 + \frac{-2a_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2}{\sigma_{ij}^2}, \\ f_{ij}(x_{sj}) = 0, \text{ иначе;} \end{cases} \\
 б) \quad & \begin{cases} f_{ij}(x_{sj}) = \frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2}, \\ f_{ij}(x_{sj}) = 0, \text{ иначе;} \end{cases} \\
 в) \quad & \begin{cases} f_{ij}(x_{sj}) = k_1 \left(\frac{1}{\sigma_{ij}^2} x_{sj}^2 + \frac{-2a_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2}{\sigma_{ij}^2} \right) + \\ + k_2 \left(\frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2} \right), \\ f_{ij}(x_{sj}) = 0, \text{ иначе;} \end{cases}
 \end{aligned}$$

$$2) \left\{ \begin{array}{l} f_{ij}(x_{sj}) = k_1 \left(\frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2} \right) + \\ \quad + k_2 \left(\frac{1}{\sigma_{ij}^2} x_{sj}^2 + \frac{-2a_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2}{\sigma_{ij}^2} \right), \\ f_{ij}(x_{sj}) = 0, \text{ иначе.} \end{array} \right.$$

Слой фаззификации. Каждый узел в этом слое генерирует сигнал, соответствующий степени выполнения нечеткого правила для выборки. Этот сигнал можно назвать степенью возбуждения нечеткого правила в отношении объекта, подлежащего классификации. Степень возбуждения по этому правилу заключается в следующем:

$$\beta_{is} = \prod_{j=1}^N f_{ij}(x_{sj}), \quad (1)$$

где N – количество классификационных признаков.

Слой дефаззификации. В этом слое вычисляются веса выходов; каждое правило влияет на каждый класс в соответствии с их весами. Если правило управляет определенной областью класса, вес между этим выходом правила и конкретным классом будет больше, чем другие веса. В противном случае веса классов малы. Вес выхода для s -й выборки, относящейся к k -му классу, вычисляется следующим образом:

$$o_{sk} = \sum_{i=1}^U \beta_{is} w_{ik}, \quad (2)$$

где w_{ik} обозначает степень принадлежности к k -му классу, который контролируется i -м правилом и U представляет собой количество правил; o_{sk} обозначает весовой выход для s -й выборки, принадлежащей k -му классу.

Слой нормализации. Выходы сети должны быть нормализованы, так как суммирование весов может быть больше чем 1, в некоторых случаях

$$h_{sk} = \frac{o_{sk}}{\sum_{l=1}^K o_{sl}} = \frac{o_{sk}}{\delta_s}, \quad \delta_s = \sum_{l=1}^K o_{sl}, \quad (3)$$

где h_{sk} представляет нормированное значение выхода для s -й выборки, которая относится к k -му классу, и K – количество классов. После этого метка класса s -й выборки определяется максимальным значением как

$$C_s = \max_{k=1,2,\dots,K} \{h_{sk}\}, \quad (4)$$

где C_s обозначает метку класса для s -й выборки.

Антецедентные параметры сети $\theta = \{S_{M \times N}, C_{M \times N}, W_{M \times K}\}$ [17] могут быть адаптированы любым методом оптимизации S и C – матриц, содержащих значения σ и μ соответственно; W представляет весовую матрицу соединений от слоя фаззификации до слоя дефаззификации; M, N, K – количество правил, признаков и классов соответственно; метод кластеризации К-средних используются для получения исходных параметров матриц S и C . В данной работе метод масштабированного сопряженного градиента (SCG) используется для адаптации параметров сети. SCG – это алгоритм сопряженного градиента второго порядка [16], который минимизирует целевые функции нескольких переменных. Теоретическая основа этого алгоритма была доказана Моллером. Этот алгоритм относится к методам первого порядка, использующих первые производные. SCG использует механизм масштабирования размера шага, который позволяет избежать трудоемкого поиска линий на каждую итерацию обучения, что повышает его быстродействие. Основанный на работе Моллера [8], метод SCG показывает суперлинейную конвергенцию [18] при решении многих задач.

Параметр w_{ik} также можно адаптировать с помощью метода SCG. Однако при обучении w_{ik} может быть больше 1. В этих случаях значения весов могут быть потеряны для кластеров из этого же класса. По этой причине либо w_{ik} должен быть ограничен, либо w_{ik} определяется из отношения числа выборки k -го класса в i -й области нечеткого правила относительно общего числа выборок k -го класса:

$$w_{ik} = \frac{D_i}{D_k}, \quad (5)$$

где D_i – число выборок k -го класса, которые принадлежат к i -й нечеткой области правил и D_k – число всех выборок k -го класса.

Функция стоимости, используемая в методе SCG, определяется из наименьших средних квадратов разностной цели и вычисленного значения класса

$$E = \frac{1}{N} \sum_{s=1}^N E_s, \quad (6)$$

$$E_s = \frac{1}{2} \sum_{k=1}^K (t_{sk} - h_{sk})^2, \quad (7)$$

где N – количество выборок, t_{sk} и h_{sk} являются целевыми и вычисленными значениями s -й выборки, принадлежащей k -му классу соответственно. Если s -я выборка принадлежит k -му классу, целевое значение t_{sk} устанавливается равным 1 и в противном случае равно 0.

Частную производную от E , относящуюся к μ_{ij} , можно вычислить так:

$$\frac{\partial E}{\partial \mu_{ij}} = \sum_{s=1}^N \frac{\partial E}{\partial E_s} \left(\sum_{k=1}^K \sum_{l=1}^K \frac{\partial E_s}{\partial h_{sk}} \frac{\partial h_{sk}}{\partial o_{sk}} \frac{\partial o_{sk}}{\partial \beta_{is}} \frac{\partial \beta_{is}}{\partial f_{ijs}} \frac{\partial f_{ijs}}{\partial \mu_{ij}} \right). \quad (8)$$

Для функции активации вида σ

$$\begin{cases} f_{ij}(x_{sj}) = \frac{-1}{\sigma_{ij}^2} x_{sj}^2 + \frac{2\mu_{ij}}{\sigma_{ij}^2} x_{sj} + \frac{a_{ij}^2 - 2\mu_{ij}a_{ij}}{\sigma_{ij}^2}, \\ f_{ij}(x_{sj}) = 0, \text{ иначе,} \end{cases} \quad (9)$$

где $a = (\mu - \sigma)$, если $\mu - \sigma \leq x \leq \mu$, и $a = (\mu + \sigma)$, если $\mu \leq x \leq (\mu + \sigma)$; $f_{ij}(x_{sj})$ представляет собой степень принадлежности i -го правила и j -й признак; x_{sj} представляет собой s -ю выборку и j -й признак входной матрицы.

Из (1)–(9) получим

$$\frac{\partial E}{\partial \mu_{ij}} = \frac{1}{N} \sum_{s=1}^N \sum_{k=1}^K \left((h_{sk} - t_{sk}) \left(\frac{1 - h_{sk}}{\delta_s} \right) w_{ik} \beta_{is} \frac{2x_{sj} - 2\mu_{ij}}{-x_{sj}^2 + 2\mu_{ij}x_{sj} + \sigma_{ij}^2 - \mu_{ij}^2} \right),$$

$$\frac{\partial E}{\partial \sigma_{ij}} = \frac{1}{N} \sum_{s=1}^N \sum_{k=1}^K \left((h_{sk} - t_{sk}) \left(\frac{1 - h_{sk}}{\delta_s} \right) w_{ik} \beta_{is} \frac{2(x_{sj} - 2\mu_{ij})^2}{\sigma_{ij}(-x_{sj}^2 + 2\mu_{ij}x_{sj} + \sigma_{ij}^2 - \mu_{ij}^2)} \right).$$

Аналогично получаются частные производные от E , относящиеся к μ_{ij} для остальных ФА.

ННС обучается методом оптимизации SCG с использованием частных производных E -отношения к приведенным выше параметрам.

3. ТЕСТИРОВАНИЕ ННК

Разработанный выше ННК с нашими функциями активации и алгоритмом SCG исследован на задаче классификации набора данных ирисы Фишера для нахождения отображения между четырьмя входными переменными (длина наружной доли околоцветника, ширина наружной, длина внутренней и ширина внутренней) и тремя классами (Setosa, Versicolor и Virginica) [11].

В наборе данных имеется 150 образцов, 75 из них используются как данные для обучения ННК, а остальные 75 – для тестирования ННК, который обучается на наборе учебных данных за 100 эпох. Затем скорректированная сеть ННК исследуется на наборе данных тестирования. Желательные и рассчитанные выходы соответствуют всем 75 тестовым данным.

Эксперимент с ННК по классификации состояний объекта проведен на наборе данных шкалы баланса, который получен из UC Irvine Machine

Learning Repository [19] и был сгенерирован для моделирования результатов психологических экспериментов, проведенных Зиглером [12]. Примеры в наборе данных соответствуют одному из трех классов: шкала справа, шкала слева или шкала сбалансирована. Набор данных состоит из 625 примеров и пяти атрибутов: имя класса (слева, баланс, справа), левый вес (1, 2, 3, 4, 5), левое расстояние (1, 2, 3, 4, 5), правый вес (1, 2, 3, 4, 5) и правое расстояние (1, 2, 3, 4, 5). Например, элемент набора данных «2, 5, 2, 1» (левый вес = 2, левое расстояние = 5, правый вес = 2, правое расстояние = 1) следует классифицировать по классу «слева». Для обучения используем 313 образцов, а для тестирования – 312.

Результаты тестирования ННК представлены в табл. 1. Эффективность работы ННК при идентификации теста будем оценивать процентным отношением между количеством правильных результатов тестирования и общим количеством данных тестирования. Чем больше процентное отношение, тем выше эффективность работы ННК на этом тесте. Данное исследование характеризует эффективность работы ННК, т. е. эффективность классификации. Среднеквадратическая ошибка (Root Mean Square Error – RMSE) характеризует точность результатов экспериментов с ННК. Чем меньше эта ошибка, тем выше точность работы ННК на этом тесте. То есть это исследование характеризует точность классификации.

Таблица 1

Table 1

Результаты классификации набора данных ННК с различными ФА

Results of classification of NFC data set with different AFs

ФА	Эффективность ННК на тесте (%)		RMSE	
	Ирисы Фишера	Шкала баланса	Ирисы Фишера	Шкала баланса
Гауссоида	96	88.78	$2.297 \cdot 10^{-7}$	0.13
Вида <i>a</i>	96	87.82	$5.031 \cdot 10^{-6}$	0.12
Вида <i>b</i>	97.33	89.10	$6.361 \cdot 10^{-4}$	0.15
Вида <i>c</i>	97.33	77.88	0.0418	0.31
Вида <i>d</i>	96	60.25	0.041	0.61

Из табл. 1 следует, что результаты классификации с использованием предлагаемых нами ФА не превосходят результат классификации с использованием гауссовой ФА (RMSE ННК с гауссовой ФА на тесте ирисы Фишера самая малая), однако использование ННК с ФА вида *a* и *b* вполне приемлемо на разных тестах. Эффективность ННК с ФА вида *b* и *c* на тесте ирисы Фишера самая высокая.

Проведем сравнительный анализ результатов классификации ННК с предлагаемыми ФА и классификации с использованием радиальной базисной нейронной сети (РБНС), являющейся популярной альтернативой известному многослойному персептрону, поскольку имеет простую структуру и более высокую скорость обучения [20]. Также РБНС использует радиальные базис-

ные функции в качестве функций активации. В табл. 2 представлены результаты эксперимента по классификации предлагаемым ННК и РБНС для выбранных тестов ирисы Фишера и шкалы баланса.

Таблица 2

Table 2

Результаты классификации ННК и РБНС

Results of classification of NFCs and RMSE

Метод использования	Эффективность на тестах (%)	
	Ирисы Фишера	Шкала баланса
ННК	96 – 97.33	60.25 – 89.10
РБНС	65.33	54.91

Из табл. 2 следует, что предлагаемый ННК эффективнее РБНС в решаемой задаче классификации.

На рис. 3 показана динамика RMSE ННК на тесте ирисы Фишера в сравнении с учебными эпохами.

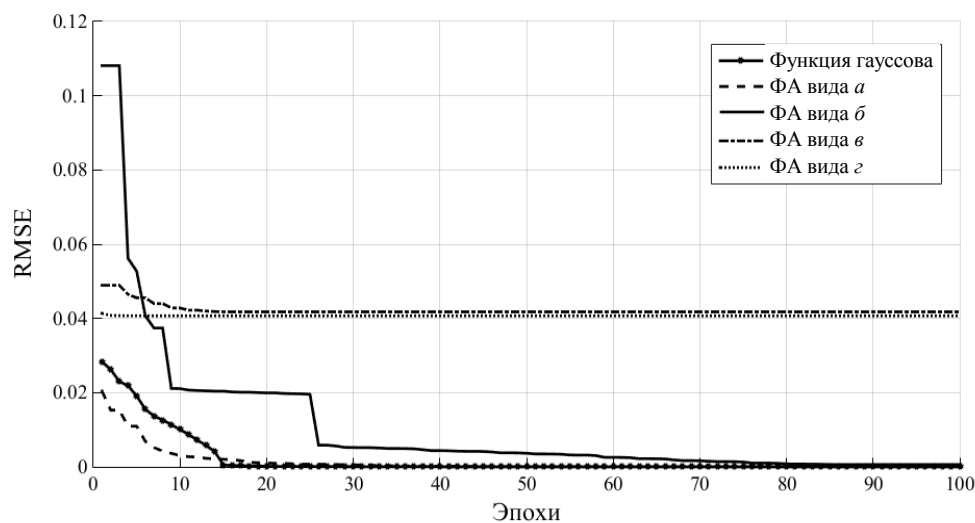


Рис. 3. RMSE ННК на тесте ирисы Фишера

Fig. 3. RMSE NFC in the Fisher's Iris test

Из рис. 3 следует, что обучение ННК с ФА α и γ на основе алгоритма SVG неэффективно. RMSE (точность работы ННК с этими ФА) почти не изменяются и слишком велики, поэтому для ННК этого типа необходим поиск других методов обучения. Относительно обучения ННК с ФА α и β на основе алгоритма SVG следует сделать более оптимистичный вывод и утверждать, что ННК данных типов эффективно решают задачи классификации объектов и их состояний.

ЗАКЛЮЧЕНИЕ

Выполнена разработка и исследование ННК, использующих четыре вида нечетких ФА, для решения задач классификации биологических объектов и их состояний. Для обучения ННК применен алгоритм масштабированного сопряженного градиента (SCG). Тестирование разработанных ННК на наборах данных ирисы Фишера и шкалы баланса, их сравнительный анализ с известным ННК с гауссовой ФА и РБНС показали конкурентоспособность двух из четырех разработанных ННК по эффективности и точности классификации. Для двух неконкурентоспособных ННК необходим поиск других методов обучения и, в частности, возможно применение метода Нелдера–Мида [21], который успешно зарекомендовал себя при обучении НС, входящей в структуру системы автоматического управления.

СПИСОК ЛИТЕРАТУРЫ

1. Галушкин А.И. Нейронные сети: основы теории. – М.: Горячая линия-Телеком, 2010. – 496 с.
2. Рассел С., Норвиг П. Искусственный интеллект: современный подход. – М.: Вильямс, 2006. – 1408 с.
3. Кориков А.М., Павлов С.Н. Теория систем и системный анализ: учебное пособие. – М.: Инфра-М, 2014. – 288 с.
4. Рыбина Г.В. Основы построения интеллектуальных систем: учебное пособие. – М.: Финансы и статистика: Инфра-М, 2010. – 432 с.
5. Браверман Э.М., Мучник И.Б. Структурные методы обработки эмпирических данных. – М.: Наука, 1983. – 464 с.
6. Nguyen A.T., Korikov A.M. Models of neural networks with fuzzy activation functions [Electronic resource] // IOP Conference Series: Materials Science and Engineering. – 2017. – Vol. 177. – URL: <http://iopscience.iop.org/article/10.1088/1757-899X/177/1/012031> (accessed: 13.09.2018).
7. Nguyen A.T., Korikov A.M. Neural network model with fuzzy activation functions for time series predictions // Доклады ТУСУР. – 2016. – Т. 19, № 4. – С. 50–52.
8. Moller M.F. A scaled conjugate gradient algorithm for fast supervised learning // Neural Networks. – 1993. – Vol. 6, N 4. – P. 525–533.
9. Jang J.S.R., Sun C.T., Mizutani E. Neuro-fuzzy and soft computing. – Upper Saddle River, NJ: Prentice Hall, 1997. – P. 110–123.
10. Ata R., Kocyigit Y. An adaptive neuro-fuzzy inference system approach for prediction of tip speed ratio in wind turbines // Expert Systems with Applications. – 2010. – Vol. 37, N 7. – P. 5454–5460.
11. Fisher R.A. The use of multiple measurements in taxonomic problems // Annals of Eugenics. – 1936. – Vol. 7. – P. 179–188.
12. Siegler R.S. Three aspects of cognitive development // Cognitive Psychology. – 1976. – Vol. 8, N 4. – P. 481–520.
13. Attaway S. MATLAB: a practical introduction to programming and problem solving. – 3rd ed. – Burlington, MA: Elsevier, 2013. – 539 p.
14. Chakraborty D., Pal N.R. A neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification // IEEE Transactions on Neural Networks. – 2004. – Vol. 15, N 1. – P. 110–123.
15. Sun C.T., Jang J.S. A neuro-fuzzy classifier and its applications // Proceedings of the IEEE International Conference on Fuzzy System. – San Francisco, 1993. – P. 94–98.
16. Методы решения задачи минимизации квадратичной функции. Проблемы сходимости: методические указания / сост. К.В. Григорьева; С.-Петербург. гос. архит.-строит. ун-т. – СПб., 2009. – 36 с.

17. Abe S. Support vector machines for pattern classification. – NewYork, NY: Springer, 2010. – 473 p.
18. Senning J.R. Computing and estimating the rate of convergence [Electronic resource] / Gordon College, Department of Mathematics and Computer Science. – URL: <http://www.math-cs.gordon.edu/courses/ma342/handouts/rate.pdf> (accessed: 13.09.2018).
19. Siegler R.S. Balance scale weight & distance database [Electronic resource] // UCI Machine Learning Repository. – URL: <http://archive.ics.uci.edu/ml/datasets/balance+scale> (accessed: 13.09.2018).
20. Using radial basis function networks for function approximation and classification / Y. Wu, H. Wang, B. Zhang, K.L. Du // ISRN Applied Mathematics. – 2012. – Vol. 2012. – Art. 324194. – 34 p.
21. Игумнов И.В., Куцый Н.Н. Метод Нелдера – Мида в обучении искусственной нейронной сети, входящей в состав ШИМ-элемента // Научный вестник НГТУ. – 2017. – № 3. – С. 22–30.

Кориков Анатолий Михайлович, доктор технических наук, профессор, заслуженный деятель науки РФ, академик МАН ВШ и МАИ, заведующий кафедрой автоматизированных систем управления Томского государственного университета систем управления и радиоэлектроники (ТУСУР), профессор Национального исследовательского Томского политехнического университета (НИ ТПУ). Основное направление научных исследований – автоматизация и оптимизация систем обработки информации и управления. Имеет более 300 научных публикаций, в том числе 10 монографий. E-mail: korikov@asu.tusur.ru

Нгуен Ань Ту, аспирант отделения автоматизации и робототехники Национального исследовательского Томского политехнического университета (НИ ТПУ). Основное направление научных исследований – нейронные нечеткие сети. Имеет 10 научных публикаций. E-mail: nguyenanhtu@tpu.ru

DOI: 10.17212/1814-1196-2018-3-73-86

A neuro-fuzzy classification of objects and their states*

A.M. KORIKOV^{1,2,a}, A.T. NGUYEN^{2,b}

¹ Tomsk State University of Control Systems and Radioelectronics, 40, Lenin Av., Tomsk, 634050, Russian Federation

² Tomsk Polytechnic University, 30, Lenin Av. Tomsk, 634050, Russian Federation

^a korikov@asu.tusur.ru ^b nguyenanhtu@tpu.ru

Abstract

In this paper, we present the development and some investigation results of a neural fuzzy network (NFN) to solve classification problems of biological objects and their states. A general approach to the problem of automatic classification of objects is studied. In this work, we use a Fisher's Iris data set for the object classification problem, and the assessment of the emotional state of objects is given on the data set of the balance scale. The NFN is a combination of fuzzy inference systems and a neural network. The structure of the NFN takes the form of a multilayer unidirectional network consisting of an input layer, a fuzzy activation function (AF) layer, a fuzzification layer, a defuzzification layer, a normalization layer and an output layer. In the fuzzy AF layer, we use fuzzy AFs that were developed and investigated in our previous studies. Herein, four fuzzy AFs corresponding to four types of NFN are investigated. The training process of the NFN is conducted by the use of the K-mean cluster method to determine

* Received 18 April 2018.

antecedent network parameters. During the training process, a scaled conjugate gradient (SCG) algorithm is used to reduce the computational effort of every learning iteration, and therefore, enhance the learning speed. The testing process is carried out with Fisher's Iris data sets and the balance scale. These data sets are classical and often used to illustrate the performance of various statistical classification algorithms. The results obtained by various types of fuzzy AF have confirmed the validity of the current NFN to solve classification problems. A comparative analysis of the current four NFN types and the NFN using the Gaussian AF from a previous paper in terms of efficiency and accuracy is performed in this work.

Keywords: neural fuzzy network, classification, fuzzy activation function, scaled conjugate gradient algorithm, Iris Fisher, scales balance, classification efficiency, classification accuracy

REFERENCES

1. Galushkin A.I. *Neironnye seti: osnovy teorii* [Neural networks: basic theory]. Moscow, Goryachaya liniya-Telekom Publ., 2010. 499 p.
2. Russel S.J., Norvig P. *Artificial intelligence: a modern approach*. 2nd ed. Upper Saddle River, NJ, Prentice Hall/Pearson Education, 2003 (Russ. ed.: Russel S., Norvig P. *Iskusstvennyi intellekt: sovremennyyi podkhod*. Moscow, Williams Publ., 2006. 1408 p.).
3. Korikov A.M., Pavlov S.N. *Teoriya sistem i sistemnyi analiz* [Theory of systems and systems analysis]. Moscow, Infra-M Publ., 2014. 288 p.
4. Rybina G.V. *Osnovy postroeniya intellektual'nykh sistem* [Fundamentals of building intelligent systems]. Moscow, Finansy i statistika Publ., Infra-M Publ., 2010. 432 p.
5. Braverman E.M., Muchnik I.B. *Strukturnye metody obrabotki empiricheskikh dannykh* [Structural methods for processing empirical data]. Moscow, Nauka Publ., 1983. 464 p.
6. Nguyen A.T., Korikov A.M. Models of neural networks with fuzzy activation functions. *IOP Conference Series: Materials Science and Engineering*, 2017, vol. 177. Available at: <http://iopscience.iop.org/article/10.1088/1757-899X/177/1/012031> (accessed 13.09.2018).
7. Nguyen A.T., Korikov A.M. Neural network model with fuzzy activation functions for time series predictions. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki – Proceedings of TUSUR*, 2016, vol. 19, no. 4, pp. 50–52.
8. Moller M.F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 1993, vol. 6, no. 4, pp. 525–533.
9. Jang J.S.R., Sun C.T., Mizutani E. *Neuro-fuzzy and soft computing*. Upper Saddle River, NJ, Prentice Hall, 1997, pp. 110–123.
10. Ata R., Kocyigit Y. An adaptive neuro-fuzzy inference system approach for prediction of tip speed ratio in wind turbines. *Expert Systems with Applications*, 2010, vol. 37, no. 7, pp. 5454–5460.
11. Fisher R.A. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 1936, vol. 7, pp. 179–188.
12. Siegler R.S. Three aspects of cognitive development. *Cognitive Psychology*, 1976, vol. 8, no. 4, pp. 481–520.
13. Attaway S. *MATLAB: a practical introduction to programming and problem solving*. 3rd ed. Burlington, MA, Elsevier, 2013. 539 p.
14. Chakraborty D., Pal N.R. A neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification. *IEEE Transactions on Neural Networks*, 2004, vol. 15, no. 1, pp. 110–123.
15. Sun C.T., Jang J.S. A neuro-fuzzy classifier and its applications. *Proceedings of the IEEE International Conference on Fuzzy System*. San Francisco, 1993, pp. 94–98.
16. Grigor'eva K.V., comp. *Metody resheniya zadachi minimizatsii kvadrachnoi funktsii. Problemy skhodimosti* [Methods for solving the problem of minimizing a quadratic function. Convergence problems]. St. Petersburg State University of Architecture and Civil Engineering. St. Petersburg, 2009. 36 p.

17. Abe S. *Support vector machines for pattern classification*. New York, NY, Springer, 2010. 473 p.
18. Senning J.R. *Computing and estimating the rate of convergence*. Gordon College, Department of Mathematics and Computer Science. Available at: <http://www.math-cs.gordon.edu/courses/ma342/handouts/rate.pdf> (accessed 13.09.2018).
19. Siegler R.S. Balance scale weight & distance database. *UCI Machine Learning Repository*. Available at: <http://archive.ics.uci.edu/ml/datasets/balance+scale> (accessed 13.09.2018).
20. Wu Y., Wang H., Zhang B., Du K.L. Using radial basis function networks for function approximation and classification. *ISRN Applied Mathematics*, 2012, vol. 2012, art. 324194. 34 p.
21. Igumnov I.V., Kutsyi N.N. Metod Nelder – Mida v obuchenii iskusstvennoi neironnoi seti, vkhodyashchei v sostav ShIM-elementa [The Nelder–Mead method in learning an artificial neural network as a part of the PWM element]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Science bulletin of the Novosibirsk state technical university*, 2017, no. 3 (68), pp. 22–30.

Для цитирования:

Кориков А.М., Нгуен А.Т. Нейро-нечеткая классификация объектов и их состояний // Научный вестник НГТУ. – 2018. – № 3 (72). – С. 73–86. – doi: 10.17212/1814-1196-2018-3-73-86.

For citation:

Korikov A.M., Nguyen A.T. Neuro-nechetkaya klassifikatsiya ob"ektov i ikh sostoyanii [A neuro-fuzzy classification of objects and their states]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Science bulletin of the Novosibirsk state technical university*, 2018, no. 3 (72), pp. 73–86. doi: 10.17212/1814-1196-2018-3-73-86.