УДК 518.9+681.51.011

ИСПОЛЬЗОВАНИЕ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ ДЛЯ ОБУЧЕНИЯ СЕТЕЙ ОБРАТНОГО РАСПРОСТРАНЕНИЯ ОШИБКИ

Муниб М.Б.[1]

[1]*Таврический национальный университет имени В.И. Вернадского, Киев, Украина (95007, г. Симферополь, пр. Вернадского, 4), e-mail: munib@mail.ua*

Обучение сетей обратного распространения ошибки происходит с использованием генетических алгоритмов. В общем случае генетические алгоритмы используются для оптимизаций во многих технических проблемах, хотя одним из самых популярных применений является использование генетических алгоритмов для оптимизации многопараметрических функций. Задача обучения нейросети это, в общем случае, задача оптимизации. Использование для обучения сетей обратного распространения ошибки генетических алгоритмов или комбинированных методов с генетическими алгоритмами позволяют выполнить обучение быстрее, однако не гарантируют оптимальных результатов. Исходя из этих соображений в статье представлена нейронная сеть обратного распространения ошибки. Приводится метод применения генетического алгоритма для обучения нейронной сети, предложена блок-схема работы алгоритма, описан процесс обучения, получен положительный результат минимизации ошибки.

Ключевые слова: генетический алгоритм, нейронные сети, процесс обучения сети, обратное распространение ошибки, слой нейронов, выходной вектор, входной вектор.

USING GENETIC ALGORITHMS FOR LEARNING NETWORKS BACK PROPAGATION OF ERROR

Munib M.B.[1]

[1]*V.I. Vernadskiy Tavricheskiy National University, Kyiv, Ukraine (95007, Simferopol, Vernadskiy Prospect 4), e-mail: munib@mail.ua*

Training networks back propagation of error comes from the use of genetic algorithms. In General, genetic algorithms are used to optimizations in many technical issues, although one of the most popular applications is to use genetic algorithms for optimization of multiparameter functions. The task of training of neural networks is, in General, the task of optimization. Use for network learning back propagation of error genetic algorithms or combined methods with genetic algorithms allow the training faster, but do not guarantee optimal results. Based on these considerations, the paper presents a backpropagation neural network errors. Provides a method of application of genetic algorithms for neural network training, offered a block diagram of the algorithm, the process of training, a positive result is obtained minimize errors.

Keywords**:** the genetic algorithm, neural network, the process of learning network, backward errors expansion, the neurons, the output vector, the input vector.

Актуальность исследования. Применение нейронных сетей для решения различных задач стало актуальным в последние несколько лет. Это, в первую очередь, задачи классификации образов, кластеризации/категоризации, аппроксимации функций, предвидения/прогнозирования, формирование памяти с содержательно зависимой адресацией, управления и оптимизации. Так, решая ту или иную задачу нейросети, прежде всего конструируют и в обязательном порядке данную сеть обучают. Обучают приему правильного, точнее адекватного, приема решения. Однако за много лет применения нейронных сетей актуальность их работы заключается в выборе топологии нейросети, алгоритма обучения, активационных функций нейронов и ряда числовых параметров для каждой практической задачи. Обучение сети может осуществляться при помощи различных

алгоритмов. Выбор необходимого подхода основывается на критериях скорости и качества обучения.

Генетические алгоритмы обучения являются достаточно мощным средством и могут с успехом применяться для широкого класса прикладных задач, включая те, которые трудно, а иногда и вовсе невозможно решить другими методами. Хотя такие алгоритмы, как и другие методы эволюционных вычислений, не гарантируют выявление глобального решения за определенно короткое время. Генетические алгоритмы не гарантируют и того, что глобальное решение будет найдено, однако они хороши для поиска адекватного решения задачи за довольно короткие сроки. Одним из преимуществ таких алгоритмов является то, что они могут применяться даже на сложных задачах, там, где не существует никаких других, эффективных методов решения. Все же на сегодняшний день вопрос исследования генетических алгоритмов для обучения сетей обратного распространения ошибки остается открытым, что и послужило выбором темы исследования.

Постановка проблемы. Останавливаясь на изучении сетей обратного распространения ошибки, на первый план выходит проблема обучения данных сетей для решения той или иной задачи, поставленной перед пользователем. Научная литература раскрывает массу трудов известных ученых, которые исследуют и применяют различные алгоритмы и методы для обучения нейросетей, говоря о достоинствах и недостатках каждого из них. Однако механизм правильного и точного обучения сетей обратного распространения ошибки не найден. Мало того, большинство предложенных методов мало эффективны и требуют достаточно длительного времени на обучение. Основываясь на этом, возникает вопрос об исследовании генетического алгоритма как основного метода для обучения сетей обратного распространения ошибки.

Анализ исследований и публикаций. В научных трудах по теме исследования освещаются различные аспекты проблематики применения генетического алгоритма для обучения сетей. В то же время заметим, что исследованием проблем обучения нейронных сетей на протяжении многих лет занимались такие отечественные и зарубежные ученые, как Р. Каллан [1], Ф. Уоссермен [2], С. Осовский [3], R. Lopez [4], О.И. Федяев, Ю.И. Соломка [5], С. Короткий [6], И.В. Заенцев [7] и другие. С момента появления известных трудов И.В. Заенцева прошли десятки лет, однако методики составления и обучения нейронных сетей продолжают развиваться достаточно бурно. Хотя на Украине практически отсутствуют исследования, которые могли бы успешно быть реализованы на практике обучения сетей. Требуют дальнейшего усовершенствования и развития методы обучения сетей обратного распространения ошибки.

Отметим, что начало изучению и усовершенствованию нейронных сетей было положено Мак-Каллоком и Питтсом. Они первыми ввели понятие «порогового логического нейрона» и описали логическую модель нейросетей [8]. Вопросом обучения нейронных сетей долгое время занимался российский ученный А.Н. Горбань [9], чему и посвятил свою работу «Обучение нейронных сетей», вышедшую в 1990 году. Этот сборник стал толчком к последующим открытиям.

Цель работы. Обучение сетей обратного распространения ошибки с использованием генетических алгоритмов. Исследование нейронной сети обратного распространения ошибки. Предложение метода применения генетического алгоритма для обучения нейронной сети, составление блок-схемы работы алгоритма, описание процесса обучения, получение положительного результата минимизации ошибки.

Изложение основного материала. Нейронная сеть обратного распространения включает в себя несколько слоев нейронов, при этом каждый нейрон *i*-го слоя связан с каждым нейроном *i+1* слоя.

Рассматривая общий случай обучения нейронных сетей, следует отметить, что этот процесс сводится к нахождению функциональной зависимости $y = f(x)$, где x - входной, а y - выходной векторы. Однако такая задача будет иметь бесконечно огромное множество решений.

Рассмотрим возможность использования генетического алгоритма для обучения сетей обратного распространения ошибки.

Анализируя исследования использования генетических алгоритмов для обучения нейросети, следует отметить, что данные алгоритмы ускоряют работу обучения, однако не являются точными. При потребности более точного решения задачи необходимо разрабатывать комбинированные методы, способные быстро закрыть пробелы сети.

Рассмотрим нейросистему, которая состоит из одного нейрона, заданного функцией активации [10]:

$$Out = {}^{1}/_{1 + e^{-(y)}} \qquad (1)$$

Структура нейронной сети представлена на рис. 1.

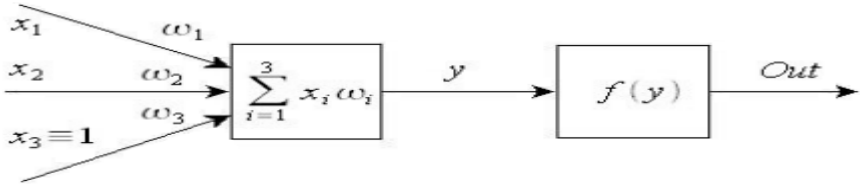

Рис. 1. Структура нейронной сети [10]

При решении задачи необходимо найти такие значения весов ω_1, ω_2, ω_3, для которых при подаче на вход сети сигналов x_1 и x_2 на ее выходе появлялись бы значение «0» или «1» в соответствии с таблицей истинности для данной логической функции. Исходя из рис. 1 очевидно, что для выбранной функции активации выход нейросети *Out* приобретает значение «0» или «1» только при $|y| \to \infty$. Исходя из этого, примем, что обучение проведено успешно, при условии, что выход сети *Out* больше 0,5 при желаемом выходе T = 1 и меньше 0,5 – при желаемом выходе T = 0 для всех случаев. Функция активации нейросети представлена на рис. 2.

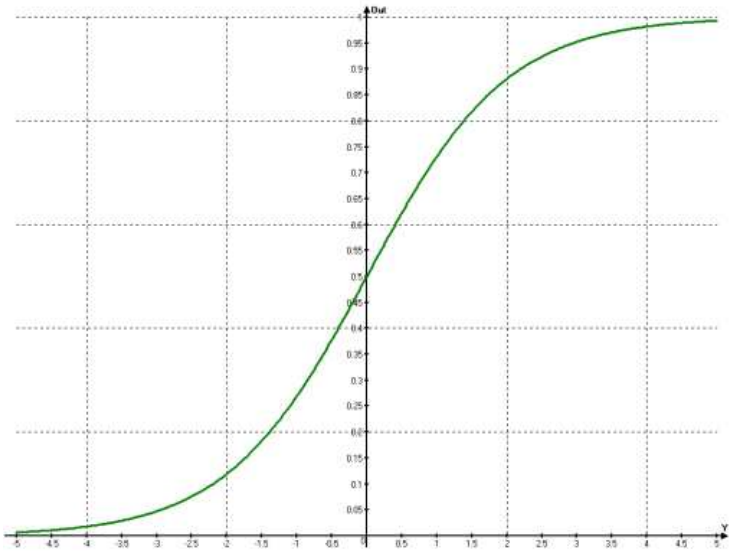

Рис. 2. Функция активации нейросети

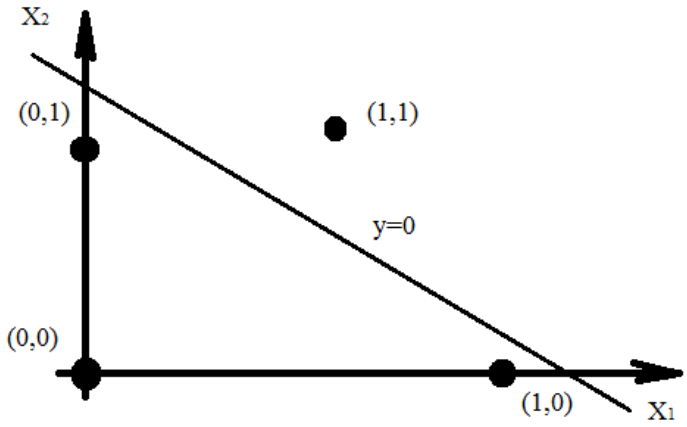

Рис. 3. Разграничение входных векторов обучающих пар на две категории

Точками на рис. 3 показано положение входных векторов обучающих пар. По условию задачи для точки (1,1) значение выхода сети *Out* должно быть больше 0,5, а для точек (0,0), (0,1) и (1,0) – меньше. При заданной функции активации выход сети *Out* , равный 0,5, когда

$$y = x_1\omega_1 + x_2\omega_2 + x_3\omega_3 = x_1\omega_1 + x_2\omega_2 + \omega_3 = 0 \qquad (2)$$

Таким образом, процесс обучения заключается в таком изменении весов ω_1, ω_2, ω_3, чтобы прямая $y=0$ разделяла точки на плоскости аналогично, как это показано на рис. 3. Все точки, что выше прямой, будут отвечать желаемому выходу $T=1$, а все, что ниже – $T=0$. Тогда значение весов ω_1, ω_2, ω_3 является решением поставленной задачи. Из рисунка видно, что таких прямых можно провести огромное множество. Следовательно, существует множество решений для данной задачи. Однако для того, чтобы определить это множество решений, необходимо воспользоваться следующими утверждениями:

1) для точек, лежащих ниже прямой y = 0, y < 0;

2) для точек, лежащих выше прямой y = 0, y > 0.

Исходя из этого, следует отметить, что если применить описанное утверждение для каждой из четырех точек, показанных на рис. 3, то можно записать следующие неравенства [7]:

$$\left\{ \begin{matrix} \omega_3 < 0 \\ \omega_2 + \omega_3 < 0; \quad \omega_1 + \omega_3 < 0; \quad \omega_1 + \omega_2 + \omega_3 > 0 \end{matrix} \right\} \rightarrow \left\{ \begin{matrix} \omega_3 < 0 \\ \omega_2 < -\omega_3; \ \omega_1 < -\omega_3; \ \omega_1 + \omega_2 > -\omega_3 \end{matrix} \right\}$$

Описанные неравенства приводят область в пространстве весов, внутри которой все точки (ω_1, ω_2, ω_3) являются решениями поставленной задачи.

Поскольку в пространстве труднее графически изобразить точки, чем на плоскости, зафиксируем одну из весов (например, ω_3) и перейдем к области решений на плоскости:

$$\left\{ \begin{matrix} \omega_3 = -1 \\ \omega_2 < 1; \ \omega_1 < 1; \ \omega_1 + \omega_2 > 1 \end{matrix} \right\}$$

На основании этого получим область решений, которая представлена на рис. 4.

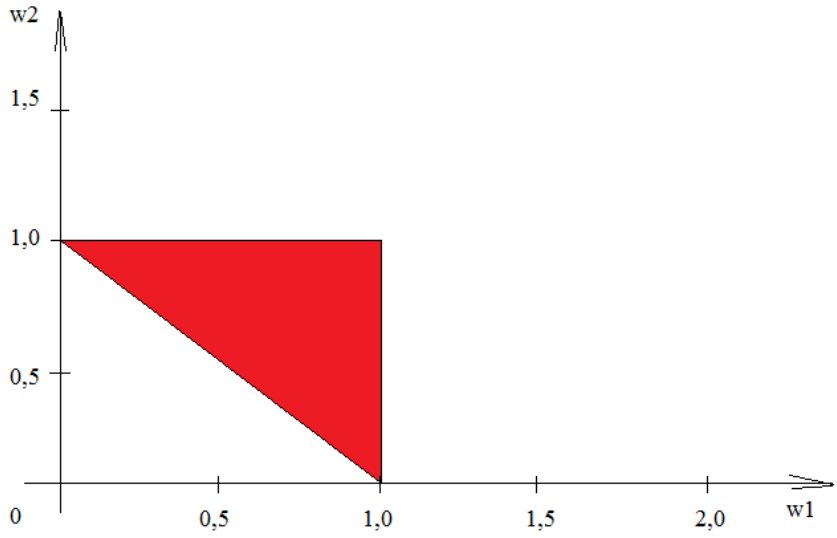

Рис. 4. Область решения задачи $\omega_3 = -1$

Таким образом, каждая точка на рис. 4 будет графическим представлением одного из вариантов решения задачи.

Далее применим генетический алгоритм для обучения нейросети. Блок-схема работы алгоритма представлена на рис. 5.

При использовании генетического алгоритма для обучения сетей обратного распространения ошибки каждая особь будет описываться двумя хромосомами: ω_1 и ω_2.

Теперь для оценки качества полученного решения зададим значение функционала погрешности в виде

$$E = {}^{1}/_{2}\,(Out - T)^2 \qquad (3)$$

где *Out* – выход сети, *T* – желаемое значение выхода. Исходя из условия, следует заметить, что количество обучающих пар равно четырем.

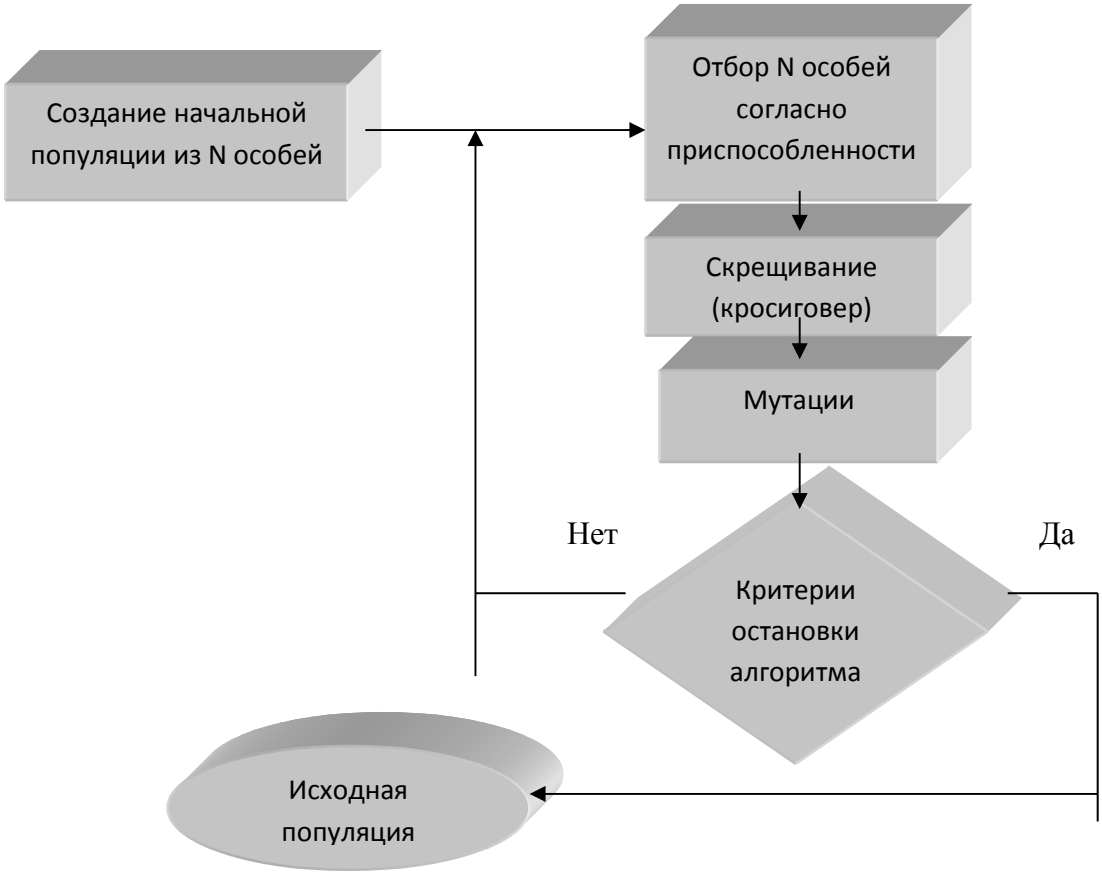

Рис. 5. Блок-схема работы генетического алгоритма

Согласно этому функционал погрешности по всем обучающим парам будет равен [7]

$$E_4 = {}^{1}/_{2} \sum_{i=1}^{4} (Out_i - T_i)^2 \qquad (4)$$

Основной задачей генетического алгоритма является минимизация погрешности E_4.

Основываясь на проведенном исследовании, максимально умным значением функционала погрешности для каждой обучающей пары будет значение E_{max}=0,125. Тогда E4 max=0,5.

Выводы

Обучение сетей обратного распространения ошибки является достаточно важным этапом в работе с нейросетями для решения задач. Исследовав нейронную сеть обратного распространения ошибки, в работе предложен метод применения генетического алгоритма для обучения нейронной сети, его суть состоит в обучении сети и получении положительного результата минимизации ошибки.

Генетический алгоритм является сочетанием градиентного и стохастического методов оптимизации функций. Регулируя вероятности кроссинговера и мутации можно достичь того, что алгоритм быстро найдет минимум. Однако этот минимум не гарантированно будет глобальным. Или же наоборот: алгоритм будет медленно спускаться к минимуму, но найдет глобальный минимум. Основным преимуществом генетического алгоритма является

относительная простота его реализации. Недостатком – большое количество переменных параметров для управления его работой.

Предложенный в работе генетический алгоритм обучения сетей обратного распространения ошибки позволяет быстро произвести обучение системы с максимально умным значением функционала погрешности для каждой обучающей пары E_{max}=0,125.

Список литературы

1. Каллан Р. Основные концепции нейронных сетей. – М. : Издательский дом «Вильямс», 2001. – 236 с.
2. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика / пер. на рус. яз., Ю.А. Зуев, В.А. Точенов, 1992 [Электронный ресурс]. – Режим доступа: http://neuroschool.narod.ru/books/nntech.html.
3. Осовский С. Нейронные сети для обработки информации / пер. с польск. И.Д. Рудинского. – М. : Финансы и статистика, 2002. – 434 с.
4. Lopez R. Neural Networks for Variational Problems in Engineering, PhD Thesis, Technical University of Catalonia, 2008 - [The electronic resource]. – Access mode: http://www.cimne.com/flood/docs/PhDThesis.pdf.
5. Федяев О.И., Соломка Ю.И. Применение генетических алгоритмов для обучения нейронных сетей. Четверта міжнародна студентська науково-практична конференція «Світ молоді – молодь світу». 15-17 квітня 2004 р. Матеріали конференції. Частина 1. – Вінниця: ВІ МАУП, 2004. – С. 85-90 [Электронный ресурс]. – Режим доступа: http://masters.donntu.edu.ua/2004/fvti/solomka/library/article2.htm.
6. Короткий С. Нейронные сети: алгоритм обратного распространения, 1996 [Электронный ресурс]. – Режим доступа: http://www.gotai.net/documents/doc-nn-003.aspx.
7. Заенцев И.В. Нейронные сети: основные модели. – Воронеж,1999. – 76 с.
8. McCulloch W.S., Pitts W. A logical calculus of the ideas immanent in nervous activity // Bulletin of Mathematical Biophysics. – 1943. – № 5. – 115-133 p.
9. Горбань А.Н. Обучение нейронных сетей. – М. : СП Параграф, 1990. – 324 с.
10. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – М. : Горячая линия – Телеком, 2001.

Рецензенты:

Иваньков И.Н., д.т.н., доцент, доцент кафедры информатики Днепропетровского национального университета имени Олеся Гончара, г. Днепропетровск.

Ковалев М.А., д.т.н., доцент, доцент кафедры информатики Днепропетровского национального университета имени Олеся Гончара, г. Днепропетровск.