

Информатика, вычислительная техника и управление

УДК 004.89

ДВУХСТАДИЙНАЯ НОРМАЛИЗАЦИЯ ВЫХОДНЫХ СИГНАЛОВ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Р.В. Гирин¹, С.П. Орлов²

¹ ООО «ИнтеллектСофт»
Россия, 443096, г. Самара, ул. Мичурина, 52

² Самарский государственный технический университет
Россия, 443100, г. Самара, ул. Молодогвардейская, 244

E-mail: romangirin@gmail.com

Рассматривается проблема интерпретации выходных сигналов искусственных нейронных сетей при применении нелинейной нормализации, вычисляемой с помощью часто применяемых на практике эвристик. Предложено использование в искусственной нейронной сети многосверточной функции активации на основе двухстадийной нормализации, выполняющей категоризацию объектов. Показано, что выходные значения, получаемые при использовании данной функции, можно интерпретировать как вероятности полной группы событий. При этом вид распределения позволяет выполнять оценку точности полученного прогноза. На основе реализованной искусственной нейронной сети, выполняющей распознавание рукописных символов от 0 до 9, произведена экспериментальная проверка предложенного подхода.

Ключевые слова: категоризация объектов, искусственные нейронные сети, нормализация.

Введение

Задача категоризации некой сущности по известному перечню ее атрибутов актуальна во многих случаях. Во многих практических задачах важно оценить не только принадлежность сущности к заданной категории, но и найти вероятность данной принадлежности [1]. В настоящей статье описано решение данной задачи применительно к оптическому распознаванию символов с использованием нейронной сети. Применение нейронной сети для категорирования тесно связано с формированием распределения вероятностей принадлежности сущности к заданному перечню категорий [1].

При решении этой задачи важную роль играет алгоритм нормализации выходных значений [2]. В статье рассмотрены теоретические аспекты различных подходов к нормализации совокупности значений, а также применение этих подходов для преобразования выходных сигналов нейронных сетей. Заложенные принципы и разработанные программные библиотеки, использованные в ходе описанных экспериментов, могут быть применены в различных технических приложениях.

Роман Викторович Гирин, ведущий инженер-программист.

Сергей Павлович Орлов (д.т.н., проф.), заведующий кафедрой «Вычислительная техника».

Двухстадийная нормализация

В общем случае под нормализацией значений понимают приведение совокупности некоторых значений к заданному диапазону [3]. Наиболее известные подходы нормализации – это линейная нормализация, нелинейная нормализация, а также нормализация по диапазону [4].

Пусть имеется совокупность каких-либо значений, далее именуемая выборкой, которую нам надо нормализовать. Линейная нормализация задается формулой (1), а нелинейная нормализация – формулой (2):

$$L(a) = \frac{a_i}{\sum_{j=1}^n a_j}, \quad (1)$$

$$L_{sm}(a) = \frac{e^{a_i - D}}{\sum_{j=1}^n e^{a_j - D}}, \quad (2)$$

где a – значение переменной из исходной выборки;

i – индекс переменной a в выборке;

n – число значений в выборке;

$$D = \max \{a_1, a_2, \dots, a_n\}.$$

В случае линейной нормализации после ее применения все нормализованные значения находятся в диапазоне от 0 до 1 и сумма этих значений равна 1. Подобное свойство позволяет трактовать полученные значения как вероятности, а всю выборку – как вероятности полной группы событий. Однако следует отметить, что линейная нормализация в случае применения к выборкам, содержащим отрицательные значения, теряет данное свойство.

После применения нелинейной нормализации все значения находятся в диапазоне от 0 до 1 и сумма всех значений, входящих в выборку, равна 1. Это справедливо и в случае присутствия в выборке отрицательных чисел.

Обладая описанными выше свойствами, этот вид нормализации в некоторых случаях прекрасно подходит. Однако у нелинейной нормализации есть одна важная особенность, из-за которой часто нормализованная выборка представлена единственным значением 1 (соответствующей максимальному значению переменной до нормализации) или 0. Подобный тип векторов называют 1-hot вектор.

Это происходит в силу следующих причин. Если не использовать корректирующую константу D , то из-за экспоненциального роста при вычислении могут быть получены очень большие значения функции. При этом их будет невозможно сохранить в используемом в программе типе данных даже при достаточно умеренных значениях a . Однако при использовании D легко заметить, что вычисление сводится к манипуляции значением $e^0 = 1$ и значениями экспоненты от отрицательных аргументов. Как известно, функция экспоненты в области отрицательных аргументов асимптотически стремится к 0.

Полученный таким образом 1-hot вектор интерпретировать как распределение вероятностей будет проблематично.

Нормализация по диапазону задается формулой следующего вида:

$$R_d(a) = \frac{a_i - \min_j \{a_j\}}{\max_j \{a_j\} - \min_j \{a_j\}}, \quad (3)$$

где a – значение переменной из заданной выборки;
 i – индекс переменной в выборке;
 j – индекс выборки;
 \min_j и \max_j – минимальное и максимальное значения в j -той выборке.

Подобный подход к нормализации приводит все значения в диапазон между 0 и 1, даже если в исходную выборку входили отрицательные числа. Однако недостаток заключается в том, что подобная нормализованная выборка не суммируется в единицу, что не позволяет интерпретировать ее как распределение вероятностей полной группы событий.

В данной статье предлагается двухстадийная нормализация, применимая в случаях, когда нужно получить выборку, представляющую собой множество значений в диапазоне от 0 до 1, сумма которых равна 1, не имеющих предрасположенности приобретать вид 1-hot вектора. Для этого предлагается последовательно применять нормализацию по диапазону и линейную нормализацию.

Двухстадийная нормализация в искусственных нейронных сетях

Искусственная нейронная сеть (ИНС) – это математическая модель, а также ее программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей, сетей нервных клеток живого организма [5]. Говоря о задачах классификации, решаемых нейронными сетями, важно отметить, что выходы сети можно интерпретировать как оценки вероятности того, что элемент принадлежит некоторому классу. В этом случае методы нормализации выходных значений нейронной сети играют ключевую роль.

Для обучения нейронной сети применяется подход, при котором выходные значения сети сопоставляются с известными значениями ожидаемых выходных сигналов. Эти значения нам известны, т. к. обучающая выборка сопровождается экземплярами наборов входных сигналов, аннотированных с ожидаемым выходным сигналом. Задание функции ошибки, выражающей различие ожидаемого и фактического выходных сигналов, служит индикатором того, насколько точно работает ИНС [8, 10].

Процесс обучения нейронной сети можно сформулировать как нахождение производной заданной целевой функции потерь и движение против градиента роста данной функции с целью повышения точности сети. В свою очередь, при нахождении градиента находятся частные производные целевой функции потерь и, по правилу цепи, производные соответствующих функций, входящих в нее [7]. Процесс обучения нейронной сети представлен на рис. 1.

Рассмотрим производную функции потерь сети, использующей двухстадийную нормализацию. При получении сигнала активации в нейронной сети он передается функции активации, которая представляет собой последовательность нормализации по диапазону (3), а затем к полученным результатам применяется линейная нормализация (1). Это позволяет получить выборку с суммой значений, равной 1, которую можно трактовать как распределение вероятностей полной группы событий.

Функция двухстадийной нормализации – это функция, принимающая вектор из n элементов и возвращающая вектор из n нормализованных элементов как результат:

$$L(R_d(a)) : \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix} \rightarrow \begin{bmatrix} L(R_d(a))_1 \\ L(R_d(a))_2 \\ \dots \\ L(R_d(a))_n \end{bmatrix}, \quad (4)$$

где $L(R_d(a))_1, L(R_d(a))_2, \dots, L(R_d(a))_n$ вычисляются по формулам (1) и (3).

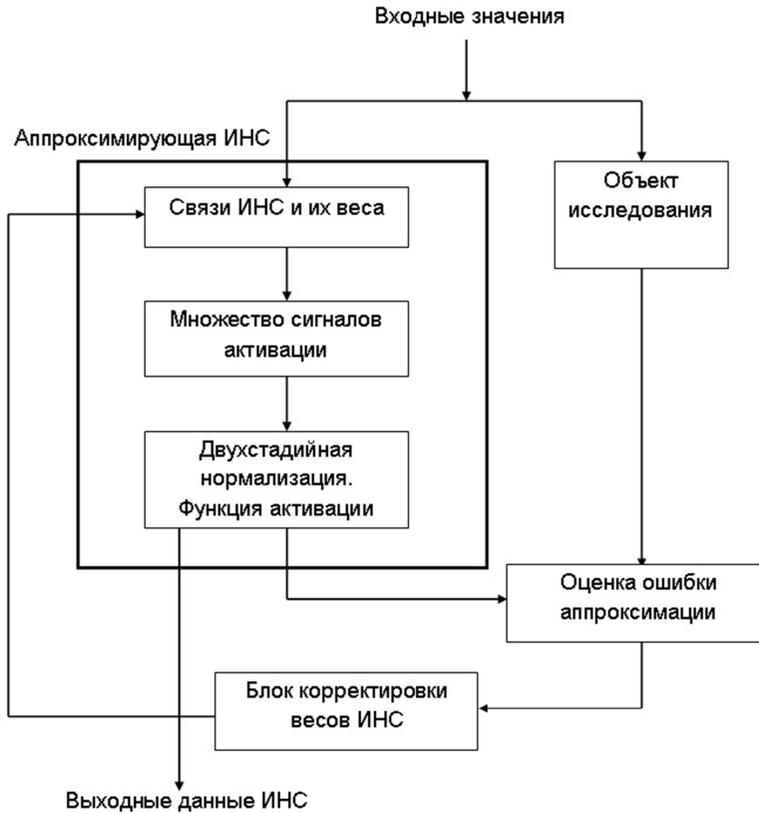


Рис. 1. Общая схема процесса обучения ИНС

В случае нахождения производной функции, принимающей вектор значений и возвращающей вектор в качестве результата, необходимо специфицировать: а) по какому выходному компоненту находится производная; б) по какой входной переменной находится частная производная. Двухстадийная нормализация заключается в последовательном применении двух функций, и тогда производная определяется по правилу цепи. Определим выражение для производной i -того выходного компонента по j -той входной переменной для нормализации по диапазону:

$$\frac{\partial R_{d_i}}{\partial a_j} = \frac{\partial \left(\frac{a_i - \min_k \{a_i\}}{\max_k \{a_i\} - \min_k \{a_i\}} \right)}{\partial a_j}, \quad (5)$$

где a – значение переменной из заданной выборки;
 i – индекс переменной в исходной выборке;

j – индекс компоненты выходного вектора, возвращаемого функцией;
 \min_k и \max_k – минимальное и максимальное значения в k -той выборке.
Получаем:

$$\frac{\partial R_{d_i}}{\partial a_j} = 1 / (\max_k \{a_i\} - \min_k \{a_i\}), \quad \forall i: i = j, \quad (6)$$

$$\frac{\partial R_{d_i}}{\partial a_j} = 0, \quad \forall i: i \neq j,$$

где использованы те же переменные, что и в выражении (5).

Для производной i -того выходного компонента по j -той входной переменной при линейной нормализации имеем:

$$\frac{\partial L_i}{\partial a_j} = \frac{-a_i + \sum_{k=1}^n a_k}{(\sum_{k=1}^n a_k)^2}, \quad \forall i: i = j, \quad (7)$$

$$\frac{\partial L_i}{\partial a_j} = \frac{-a_i}{(\sum_{k=1}^n a_k)^2}, \quad \forall i: i \neq j \quad (8)$$

Важно заметить, что в искусственной нейронной сети функции активации передают значения входных сигналов, скорректированные в зависимости от весов, что можно выразить следующим образом:

$$g_i(W_{ij}) = W_{i1}x_1 + W_{i2}x_2 + \dots + W_{it}x_t, \quad (9)$$

где W_{ij} – вес связи j -того входа ИНС и i -того нейрона;
 x – входной сигнал, поступивший в ИНС;
 t – число входных сигналов.

Вся цепочка преобразований, выполняемая в искусственной нейронной сети (обозначим ее как функцию $P(W)$), может быть выражена следующим образом:

$$P(W) = L(R_d(g(W))),$$

где функции $L(a)$, $R_d(a)$ и $g(W)$ определены в (1), (3) и (9) соответственно.

Поскольку оценка точности результатов, выдаваемых нейронной сетью, по сравнению с известным правильным ответом в ходе обучения выполняется с использованием кросс-энтропии [2], получаем следующие формулы для определения значений ее частных производных:

$$\frac{\partial E(W)}{\partial W_{ij}} = \frac{\sum_{k=1}^n a_k}{a_y} \cdot \frac{(\sum_{k=1}^n a_k) - a_y}{(\sum_{k=1}^n a_k)^2} \cdot \frac{x_j}{\max_s \{b_s\} - \min_s \{b_s\}}, \quad \forall i: i = j, \quad (10)$$

$$\frac{\partial E(W)}{\partial W_{ij}} = - \frac{x_j}{(\max_s \{b_s\} - \min_s \{b_s\}) \sum_{k=1}^n a_k}, \quad \forall i: i \neq j, \quad (11)$$

где j – индекс, нумерующий столбец матрицы весов (совпадает с индексом, нумерующим элемент входного сигнала);

i – индекс, нумерующий строку матрицы весов;

n – число компонент в векторе, возвращаемом функциями $L(a)$, $Rd(a)$ и $g(W)$;

u – индекс элемента в векторе, который соответствует известному нам результату, отмеченному в обучающем нейронную сеть примере как верный (ожидаемый) результат;

a – аргумент функции линейной нормализации;

b – аргумент функции нормализации по диапазону;

\min_s и \max_s – минимальное и максимальное значения в s -ной выборке.

Анализируя (10) и (11), легко заметить, что значения, которые корректируют веса нейронной сети, отличны от 0 не только для нейрона, соответствующего известному нам результату, который отмечен в обучающем примере как верный (ожидаемый) результат. Таким образом, выполняется корректировка всех весов в процессе обучения, а не отдельно взятых, как это имеет место в случае применения только нормализации по диапазону.

Построение нейронной сети с двухстадийной нормализацией в качестве функции активации

Для сравнения различных подходов нормализации были проведены эксперименты с ИНС, в которых в качестве функций активации применялась в одном случае нелинейная нормализация, а в другом – двухстадийная нормализация. Подобная нормализация часто применяется как завершающий слой нейронных сетей, которые выполняют категорирование.

Структура искусственных нейронных сетей, применявшихся в экспериментах, изображена на рис. 2.

Параметры сетей, использовавшихся в экспериментах, приведены в табл. 1.

Для экспериментов выбран тестовый набор (датасет) Национального института стандартизации NIST США. Содержание базы датасетов, как указано в [6], представляет собой 814 255 сегментированных рукописных цифр и букв из специально заполненных форм. Эти символы представлены в виде растровых изображений размером 128×128 пикселей и снабжены маркером, показывающим, какой именно символ изображен в виде 62 ASCII кодов, соответствующих цифрам и символам: “0” – “9”, “A” – “Z” и “a” – “z”.

Соответственно, структура ИНС, обрабатывающих данный датасет, имеет 4900 входных сигналов. Такое число входов необходимо для передачи в сеть изображения размером 70×70 . На каждый из входов подавалось значение, соответствующее значению пиксела. Все растровые изображения в базе NIST черно-белые, поэтому для передачи входного сигнала пикселы интерпретировали следующим образом: пикселу черного цвета соответствовала единица, передаваемая на соответствующий вход нейронной сети, а пикселу белого цвета соответствовал ноль. Перед началом обучения веса сетей инициализировались случайными значениями, находящимися в диапазоне от -2.0 до 2.0.

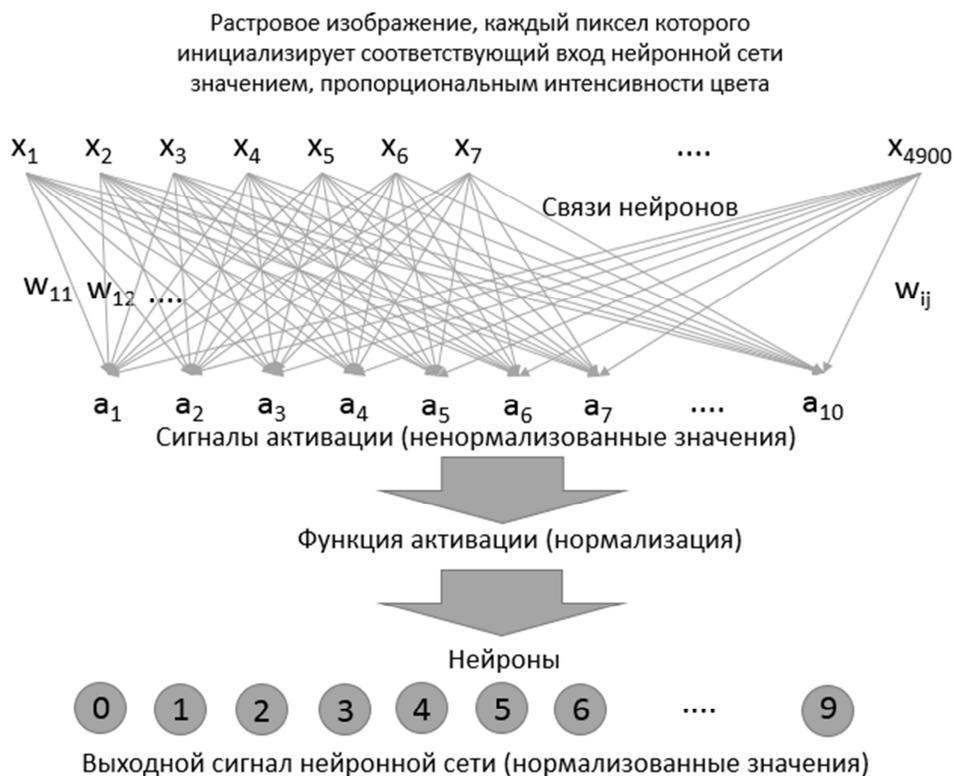


Рис. 2. Структура нейронных сетей

Таблица 1

Параметры ИНС

Параметр	Значение
Число слоев	1
Число нейронов	10
Число входов	4900
Число связей	49000

Сети обучались на части базы, которая содержит цифры. Подходы, изложенные в данной работе, применимы и для сетей, выполняющих категорирование (распознавание) не только цифр, но и всех остальных символов, представленных в базе датасетов.

Результаты экспериментов применения двухстадийной нормализации

Были построены нейронные сети, в одной из которых качестве функций активации применялась нелинейная нормализация, а в другой – двухстадийная нормализация. После обучения сети были протестированы на выборке, представленной в директории hsf_4 в базе датасетов.

Искусственные нейронные сети активно применяются для распознавания образов на изображениях [9].

Рассмотрим пример обработки изображения с цифрами. Выполним сравнение двух методов: нелинейной нормализации и двухстадийной нормализации. Распределение значений, полученных на выходе искусственной нейронной сети, в качестве функции активации, в которой применена двухстадийная нормализация, приведено в табл. 2. В этой же таблице для сравнения приведены выходные сигналы ИНС, в которой применялась нелинейная нормализация.

Оба распределения характерны для каждой из использованных в эксперименте ИНС. Как видно при сопоставлении этих распределений, в случае применения нелинейной нормализации на выходе наблюдается 1-hot вектор. При применении двухстадийной нормализации выходные значения представляют собой значения в диапазоне от 0 до 1, которые являются долей единицы.

Подобные значения можно интерпретировать как вероятность принадлежности распознаваемого символа к той или иной категории. Категория может быть оценена по наибольшему значению, при этом его абсолютная величина может служить оценкой степени точности результата.

Кроме проверки работы обученных нейронных сетей на тестовой выборке целесообразно получить визуализацию весов связей сети. Подобная визуализация может быть получена следующим образом. Если взять матрицу весов и выбрать из нее все веса связей от всех входов сети к первому нейрону (нейрону, ответственному за категоризацию изображения, на котором цифра «0»), то получатся веса связей сети, которые ИНС «выучила» для того, чтобы распознавать 0. Аналогично можно выбрать все веса, соответствующие другим категориям (в нашем случае – другим цифрам, представленным на растровых изображениях). При применении подобной группировки будет получено 10 групп (по числу категорий) весов, при этом в каждой группе будет 4900 связей. Если полученные 4900 связей изобразить на растровом изображении 70×70 пикселей, присвоив каждому пикселю значение цвета тем темнее, чем более большой вес мы имеем, то будет получена наглядная визуализация весов сети. При этом черному цвету соответствует большое значение веса, а белому – малое.

Результат такой визуализации представлен на рис. 3. Легко видеть, что искусственная нейронная сеть в процессе обучения выделила позиции пикселей, наиболее характерные для каждой из цифр. Этим позициям и соответствующим связям от данных входов искусственной нейронной сети присвоены наибольшие веса. Одновременно с этим для каждой из цифр в процессе обучения ИНС выделились позиции пикселей, наименее характерные для данной цифры. Таким позициям соответствуют наименьшие веса.

Полученное распределение значений весов и анализ его визуализации подтверждают корректность функционирования искусственной нейронной сети, в качестве функции активации в которой применена двухстадийная нормализация, при распознавании рукописных символов от 0 до 9. Однако важно отметить, что двухстадийная нормализация, примененная в качестве функции активации искусственной нейронной сети, применима для выполнения категоризации и в других предметных областях, в которых доступна репрезентативная обучающая выборка.

добные реализации нейронных сетей, выполняющих экспертные оценки, будут размещаться на сервере приложений ИНС. Это позволит клиентским устройствам со значительными ограничениями в вычислительной мощности отправлять необходимые запросы на сервер и получать результаты с искомыми оценками и значениями метрик.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Norvig P., Russell S.* Artificial Intelligence: A Modern Approach, Edition: 3rd, Pearson, 2010.
2. *Nielsen M.* Neural Networks and Deep Learning, free online book. – <http://neuralnetworksanddeeplearning.com>, 2017.
3. *Dodge Y.* The Oxford Dictionary of Statistical Terms, OUP, 2003.
4. Goodfellow I., Bengio Y., Courville A. Deep learning. – <http://www.deeplearningbook.org>, 2017.
5. *Мак-Каллок У.С., Пумтс В.* Логическое исчисление идей, относящихся к нервной активности // Автоматы / Под ред. К.Э. Шеннона и Дж. Маккарти. – М.: Изд-во иностр. лит., 1956. – С. 363–384.
6. *Grother P., Hanaoka K.* NIST Special Database 19 Handprinted Forms and Characters, 2nd Edition, National Institute of Standards and Technology, September 13, 2016.
7. *Rumelhart D.E., Hinton G.E., Williams R.J.* Learning Internal Representations by Error Propagation // Parallel Distributed Processing. Vol. 1. – Cambridge, MA: MIT Press, 1986.
8. *Комарцова Л.Г., Максимов А.В.* Нейрокомпьютеры. – М.: МГТУ им. Н.Э. Баумана, 2004.
9. *Галушкин А.И.* Синтез многослойных систем распознавания образов. – М.: Энергия, 1974.
10. *Барцев С.И., Охонин В.А.* Адаптивные сети обработки информации. – Красноярск: Ин-т физики СО АН СССР, 1986.

Статья поступила в редакцию 11 сентября 2017 г.

TWO-STAGE NORMALIZATION OF OUTPUT SIGNALS OF ARTIFICIAL NEURAL NETWORKS

R. Girin¹, S.P. Orlov²

¹ООО 'IntellectSoft'
52, Michurina st., Samara, 443096, Russian Federation

²Samara State Technical University
244, Molodogvardeyskaya st., Samara, 443100, Russian Federation

In the article the problem of interpretation output signals of artificial neural networks was considered in cases when softmax regression is used in them which calculated with heuristics often applied in practice. To use alternative multi-fold activation function in artificial neural network for objects categorizing is proposed. It is shown that output signals, that can be get in case of using the function, possible to be interpreted as probability of full group of events. Additionally, the form of distribution allows to estimate how precise the proposition is. With aim of developed artificial neural net that recognizes hand-written characters from 0 through 9 experiments were made that checks proposed approach.

Keywords: categorization, artificial neural networks, normalization of values.

*Roman V. Girin, Lead Software Engineer.
Sergey P. Orlov (Dr. Sci. (Techn.)), Professor.*