

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ВСЕРОССИЙСКАЯ ГРУППА ТЕОРИИ ИНФОРМАЦИИ ИЕЕЕ
АКАДЕМИЯ ИНФОРМАТИЗАЦИИ ОБРАЗОВАНИЯ
ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ООО «ОТКРЫТЫЕ РЕШЕНИЯ»
ОБЩЕСТВО «ЗНАНИЕ» РОССИИ
ПРИВОЛЖСКИЙ ДОМ ЗНАНИЙ

*XXII Международная
научно-техническая конференция*

**ПРОБЛЕМЫ ИНФОРМАТИКИ
В ОБРАЗОВАНИИ, УПРАВЛЕНИИ,
ЭКОНОМИКЕ И ТЕХНИКЕ**

Сборник статей

Декабрь 2022 г.

Пенза

УДК 004
ББК 32.81я43+74.263.2+65.050.2я43
П781

П781 **ПРОБЛЕМЫ ИНФОРМАТИКИ В ОБРАЗОВАНИИ,
УПРАВЛЕНИИ, ЭКОНОМИКЕ И ТЕХНИКЕ :**
сборник статей XXII Международной научно-технической
конференции. – Пенза: Приволжский Дом знаний, 2022. – 356 с.

ISBN 978-5-8356-1800-2
ISSN 2311-0406

Под редакцией *В.И. Горбаченко*, доктора технических наук,
профессора;
В.В. Дрождина, кандидата технических наук,
профессора

Информация об опубликованных статьях предоставлена в систему Рос-
сийского индекса научного цитирования (РИНЦ) по договору
№ 573-03/2014К от 18.03.2014.

ISBN 978-5-8356-1800-2
ISSN 2311-0406

© Пензенский государственный
университет, 2022
© АННМО «Приволжский Дом знаний», 2022

*XXII International
scientific and technical conference*

**PROBLEMS OF INFORMATICS
IN EDUCATION, MANAGEMENT,
ECONOMICS AND TECHNICS**

December, 2022

Penza

2. МАШИННОЕ ОБУЧЕНИЕ И НЕЙРОННЫЕ СЕТИ

УДК 681.3
ББК 32.973

ФУНКЦИИ АКТИВАЦИИ НЕЙРОННЫХ СЕТЕЙ

Л. В. Бунина, М. А. Лихачев, А. П. Титов

ACTIVATION FUNCTIONS OF NEURAL NETWORKS

L. V. Bunina, M. A. Likhachev, A. P. Titov

Аннотация. В этой статье рассмотрены функции активации нейронных сетей. Проанализированы основные характеристики и рассмотрены преимущества и недостатки. В заключение сделаны выводы о современных перспективах функций активации.

Ключевые слова: функция активации, нейрон, нейронные сети.

Abstract. This article discusses the activation functions of neural networks. The main characteristics are analyzed and the advantages and disadvantages are considered. In conclusion, conclusions are drawn about the modern prospects for activation functions.

Key words: activation function, neuron, neural networks.

В искусственных нейронных сетях (ИНС) блоки разрабатываются как свободная копия биологических нейронов в мозгу. Первый искусственный нейрон состоял из линейного порогового блока. Это было имитацией того, что искусственные нейроны не просто выводят необработанные входные данные, которые они получают. Вместо этого они выводят результат функции активации. Это поведение свойственно биологическим нейронам. Нейрон работает или нет обычно в зависимости от того, какие входные данные он получает. От моделей Персептрон до более современных архитектур глубокого обучения использовались различные функции активации. Рассмотрим самые распространенные функции активации, встречающиеся в ИНС. Обратим внимание на их свойства, а также на их производительность в общих задачах машинного обучения.

Кусочно-линейная функция

Эта функция состоит из одной из самых простых функций активации.

$$\Phi(x) = \begin{cases} 0 & x < -b, \\ x + b & -b < x < b, \\ 1 & x > b, \end{cases}$$

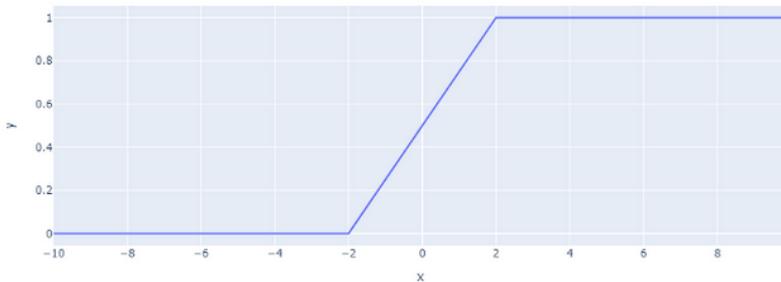


Рис. 1. Кусочно-линейная функция с $b=2$

Эта функция активации используется только как результат задачи регрессии.

Логистическая функция

Логистическая функция или логистическая кривая - самая общая сигмоидальная (S-образная) кривая. Для автоматического распознавания речи используется функция, которая определяется как:

$$\Phi(x) = \frac{1}{1 + e^{-x}}.$$

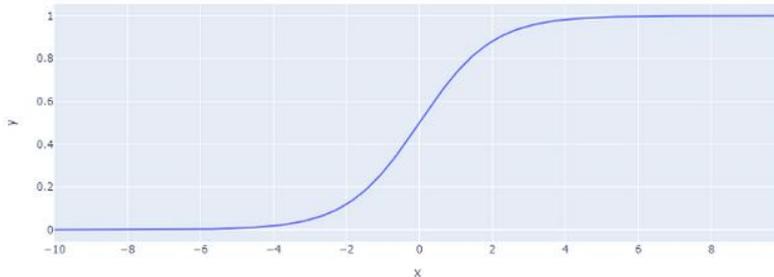


Рис. 2. Сигмовидная функция

Сигмовидная функция ограничена, что и послужило причиной ее популярности. Однако она подвержена проблеме исчезающего градиента [2], чем глубже нейронная сеть, тем менее эффективно можно обучать ее с помощью сигмоиды. Эта функция дифференцируема, что делает ее очень подходящим кандидатом для алгоритма обратного распространения ошибки.

Функция гиперболического тангенса

В начале 2000-х гиперболическая тангенсная функция заменила сигмовидную. Она находится в диапазоне от $[-1, 1]$, что делает его подходящим для задач регрессии из-за свойства центрирования нуля. Основным недостатком этой функции является насыщение. Действительно, гиперболический тангенс насыщается довольно быстро (быстрее, чем сигмовидный), что может затруднить для искусственной нейронной сети соответствующее изменение весов во время обучения.

$$\operatorname{th} x = \frac{e^{2x} - 1}{e^{2x} + 1}$$

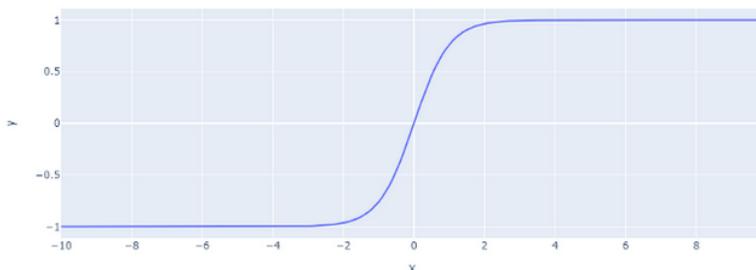


Рис. 3. Функция гиперболического тангенса

Однако стоит отметить, что эта функция часто используется в качестве функции активации для блоков скрытых слоев в рекуррентных нейронных сетях.

Разработка оптимальной функции активации.

Параметрические функции активации

Пример, использовать параметрическую функцию гиперболического тангенса [4]:

$$\Phi(x) = \frac{a(1 - \exp(-bx))}{1 + \exp(-bx)},$$

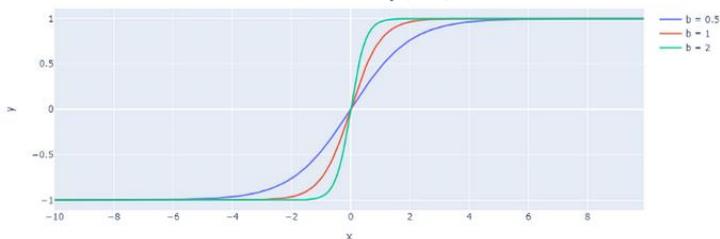


Рис. 4. Пример функции параметрического гиперболического тангенса

Преимущество параметрических функций активации заключается в том, что таким образом можно модифицировать почти любую стандартную функцию активации, добавляя параметры. Добавленные параметры усложняют вычисления, они используются в современной архитектуре глубокого обучения.

Стохастические адаптивные функции активации

Другой способ ввести адаптивные функции – использовать стохастический подход. Впервые представленный [5]. Он состоит в использовании структурированного и ограниченного шума для более быстрого обучения.

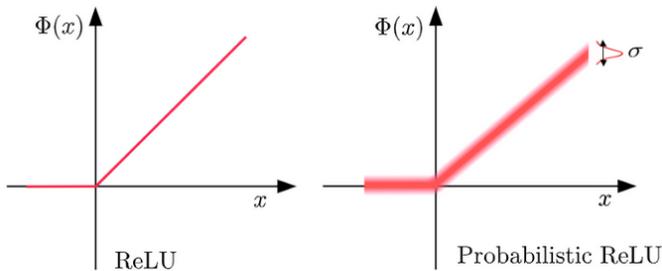


Рис. 5. Стохастический адаптивный ReLU

Это особенно полезно для функций активации с насыщением. Этот тип шумовой функции активации полезен при насыщении, поскольку шум применяется после порога, поэтому он позволяет функции превысить его. Как и в случае с параметрической функцией активации, эти функции превосходят фиксированную функцию активации в большинстве задач машинного обучения.

Одной из основных задач является получение быстрой сходимости алгоритма обучения. Этого можно добиться с помощью адаптивных функций, несмотря на их высокую вычислительную стоимость. Их более высокие показатели наблюдались как в задачах машинного обучения классификации, так и в задачах регрессионного машинного обучения. Однако доказательством того, что лучшая функция активации до сих пор не найдена, является то, что исследователи в настоящее время изучают квантованную адаптивную функцию активации, чтобы снизить вычислительные затраты при сохранении быстрой сходимости адаптивной функции.

Библиографический список

1. Maas, Andrew L., Awni Y. Hannun, and Andrew Y. Ng. "Rectifier nonlinearities improve neural network acoustic models." *Proc. icml*. Vol. 30. №1. 2013.
2. Корепанов В.Д., Кулагин В.П., Халабия Р.Ф. Использование нейронных сетей при модификации генетического алгоритма // Информационные технологии. Т. 24, №12. 2018. С. 799-804.
3. Солнцев Д.С., Халабия Р.Ф. Классификация изображения документа на группы письменности // Сборник статей по материалам научно-технической конференции молодых ученых. Волгоград: Сириус, 2021.

Бунина Людмила Владимировна
Лихачев Матвей Алексеевич
Титов Андрей Петрович
МИРЭА – Российский
технологический университет,
г. Москва, Россия

Bunina L. V.
Likhachev M. A.
Titov A. P.
MIREA – Russian Technological
University,
Moscow, Russia

УДК 004.032.26

РЕАЛИЗАЦИЯ В MATLAB ОСНОВАННЫХ НА ФИЗИКЕ НЕЙРОННЫХ СЕТЕЙ

В.И. Горбаченко, З.А. Карманова

IMPLEMENTATION IN MATLAB OF PHYSICS-INFORMED NEURAL NETWORKS

V. I. Gorbachenko, Z. A. Karmanova

Аннотация. Показана возможность реализации в MATLAB основанных на физике нейронных сетей. Приведена постановка задачи и функция потерь решения дифференциальных уравнений в частных производных с помощью основанных на физике нейронных сетей. Разработан алгоритм решения уравнений в частных производных в MATLAB с помощью основанных на физике нейронных сетей. С помощью основанных на физике нейронных сетей решена в MATLAB модельная задача, описываемая уравнением Пуассона.