

DOI: <https://doi.org/10.60797/IRJ.2024.144.174>

СРАВНЕНИЕ ФУНКЦИЙ АКТИВАЦИИ SOFTMAX И СИГМОИДА В КОНТЕКСТЕ МНОГОКЛАССОВОЙ
КЛАССИФИКАЦИИ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

Научная статья

Боровик И.Г.^{1,*}, Нгуен К.Т.²

¹ ORCID : 0000-0002-6961-6212;

² ORCID : 0009-0000-5608-677X;

^{1,2} Московский государственный технический университет им. Н.Э. Баумана, Москва, Российская Федерация

* Корреспондирующий автор (big[at]bmstu.ru)

Аннотация

Данная статья исследует важную проблему, связанную с применением функции softmax в задачах многоклассовой классификации объектов. В статье описывается структура и работа искусственного нейрона, а также принцип действия функций активации на примере сумматора. В частности, рассматривается ситуация, когда входные данные не соответствуют ни одному из классов в наборе данных, что может привести к неправильным выводам модели. Это ставит под сомнение точность классификации и требует системного контроля и обработки таких случаев. В статье предлагается альтернативный подход с использованием сигмоидальной функции активации и установлением порогового значения для определения принадлежности объекта к классам классификации. Этот подход позволяет более точно оценивать результаты классификации и избегать ложных выводов.

Ключевые слова: компьютерное зрение, нейронные сети, функция активации, softmax, сигмоидальная функция активации, классификация объектов, распознавание, обработка изображений.

A COMPARISON OF SOFTMAX AND SIGMOID ACTIVATION FUNCTIONS IN THE CONTEXT OF MULTI-
CLASS OBJECT RECOGNITION CLASSIFICATION

Research article

Borovik I.G.^{1,*}, Nguyen Q.T.²

¹ ORCID : 0000-0002-6961-6212;

² ORCID : 0009-0000-5608-677X;

^{1,2} Bauman Moscow State Technical University, Moscow, Russian Federation

* Corresponding author (big[at]bmstu.ru)

Abstract

This article studies an important problem related to the application of softmax function in multiclass object classification tasks. The paper describes the structure and operation of an artificial neuron, as well as the principle of activation functions on the example of an adder. In particular, it considers the situation when input data does not correspond to any of the classes in the dataset, which can lead to incorrect model outputs. This calls into question the accuracy of classification and requires systematic control and handling of such cases. This work proposes an alternative approach using a sigmoidal activation function and setting a threshold value to determine whether an object belongs to the classification classes. The approach allows for a more accurate evaluation of classification results and avoiding false conclusions.

Keywords: computer vision, neural networks, activation function, softmax, sigmoidal activation function, object classification, recognition, image processing.

Введение

В создании искусственных нейронных сетей одним из ключевых аспектов является выбор функции активации. Эта функция определяет поведение нейрона и формирует выходной сигнал в зависимости от входных данных. Среди различных функций активации функции softmax и сигмоиды играют особенно важную роль в задачах классификации.

В данном исследовании мы сосредоточимся на проблеме, которая возникает при использовании функции softmax в задачах многоклассовой классификации, когда входные данные не соответствуют ни одному из классов в наборе данных. Это может привести к неправильным выводам и потребности в контроле таких ситуаций.

Далее мы предложим альтернативный подход, в котором используется функция активации сигмоиды и устанавливается пороговое значение для определения принадлежности объекта к классам классификации. Этот подход позволяет улучшить точность классификации и избежать ложных результатов.

Основы функций активации и математические уравнения функции softmax и сигмоиды

Одним из важных факторов при создании искусственных нейронных сетей вообще и нейронных сетей распознавания объектов в частности является выбор функции активации. От выбранной функции активации зависят функциональные возможности нейронной сети и выбор метода для её обучения. Функция активации описывает выходное значение нейрона в зависимости от взвешенной суммы его входов и порогового значения срабатывания [1]. На рисунке 1 представлена структурная схема искусственного нейрона.

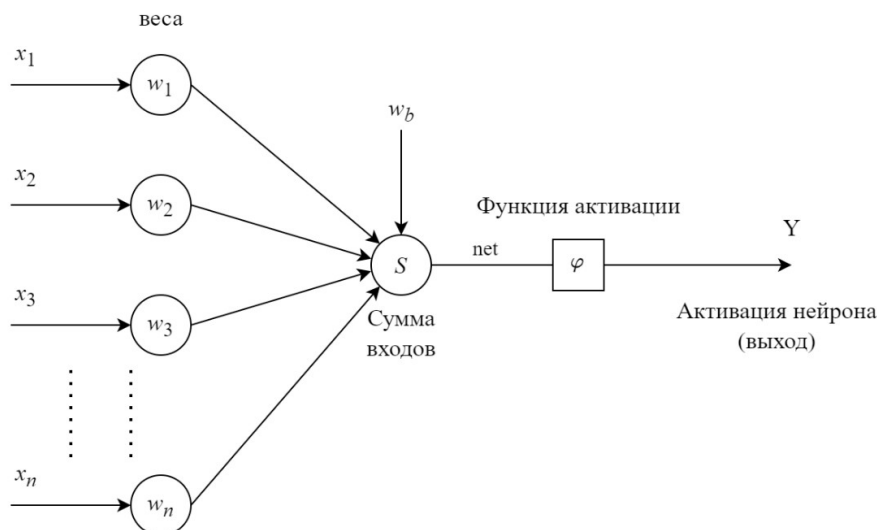


Рисунок 1 - Структурная схема искусственного нейрона
DOI: <https://doi.org/10.60797/IRJ.2024.144.174.1>

Нейронный сумматор осуществляет алгебраическое суммирование взвешенных входных сигналов. Математическое выражение для вычисления выходного сигнала сумматора S имеет следующий вид:

$$S = \sum_{i=1}^n x_i * w_i ;$$

где: x_i – выходной сигнал i -го нейрона;

w_i – вес i -го нейрона;

n – количество нейронов, входящих в обрабатываемый нейрон.

Сумматор генерирует выходной сигнал net , который затем преобразуется функцией активации. Формула для вычисления net :

$$net = \sum_{i=1}^n x_i * w_i + w_b ;$$

где w_b – сигнал смещения (bias), отображающий функцию предельного значения.

Функция активации $f(S)$ обрабатывает выходной сигнал сумматора net и формирует нейронный сигнал выхода Y .

В задачах распознавания объектов на изображениях доминируют сверточные нейронные сети (СНС). СНС состоит из сверточных слоев и слоев пулинга, последовательно соединенных друг с другом. Далее эти слои подключаются к полностью связанному слою, который формирует выходные данные. Перед формированием выходных данных используются функции активации. Наиболее распространенные функции активации: Функция softmax и сигмоидальная функция.

Сигмоидальная функция

Сигмоидальная функция активации – это нелинейная функция, которая преобразует входное значение в диапазоне от отрицательной бесконечности до положительной бесконечности в значение от 0 до 1. Эта функция активации часто используется в нейронных сетях для задач бинарной классификации [2].

Сигмоидальная функция активации представлена формулой [3]:

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

Параметр α влияет на крутизну функции. Чем он больше, тем более круче функция. Данным параметром можно задать диапазон входных значений, которые не будут прилегать к минимуму или максимуму функции. Графически сигмоидальная функция активации выглядит как S-образная кривая, которая монотонно возрастает и имеет асимптоты на 0 и 1. В частности, если $x > 0$, то $f(x) > 0.5$, а если $x < 0$, то $f(x) < 0.5$. Значение 0.5 достигается при $x = 0$ (см. рис. 2).

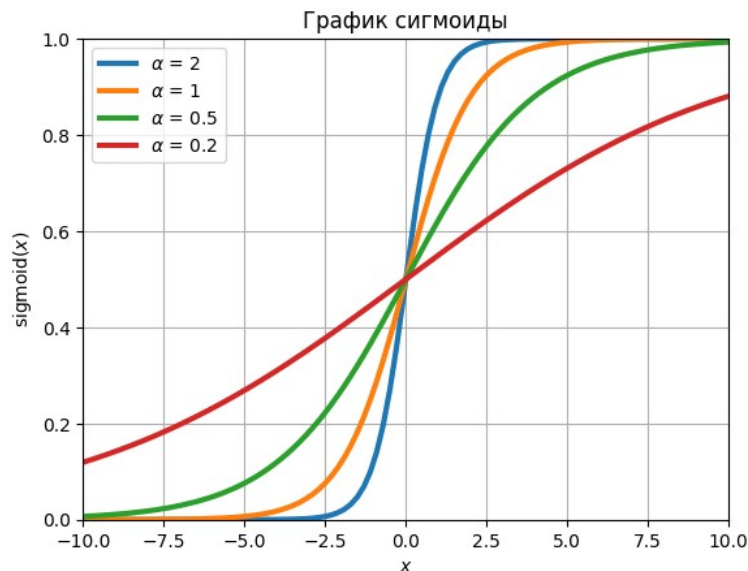


Рисунок 2 - График сигмоидальной функции активации
DOI: <https://doi.org/10.60797/IRJ.2024.144.174.2>

Функция Softmax

Функция Softmax используется для преобразования вектора значений в вероятностное распределение, которое суммируется до 1 [4]. Она особенно полезна в многоклассовой классификации, где необходимо определить вероятности для каждого класса.

Формула функции Softmax выглядит следующим образом:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} ;$$

где z_i – это элемент входного вектора, а k – это общее число элементов в векторе.

График функции Softmax представляет собой гладкую кривую, начинающуюся от 0 и заканчивающуюся на 1, что соответствует сумме вероятностей всех элементов вектора. Кривая функции Softmax имеет свойство, что вероятность любого элемента вектора увеличивается, если значения других элементов уменьшаются, что позволяет использовать эту функцию для многоклассовой классификации [5] (см. рис. 3).

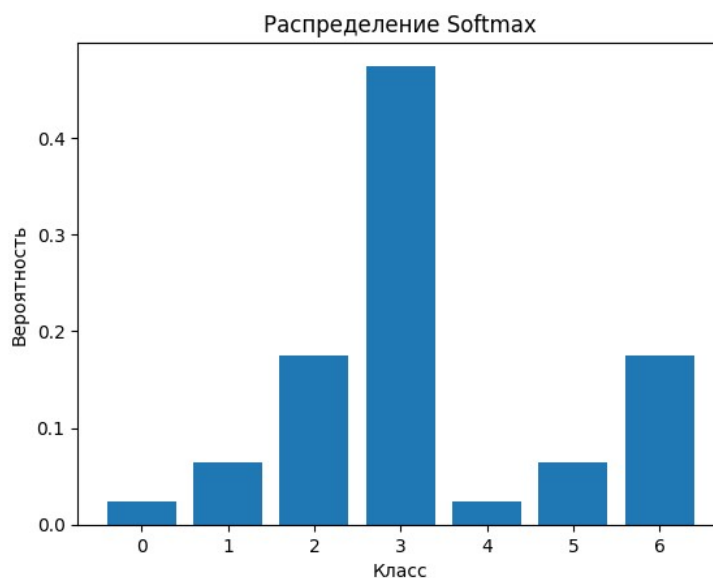


Рисунок 3 - График функции softmax, соответствующий входному вектору $z = [1, 2, 3, 4, 1, 2, 3]$
DOI: <https://doi.org/10.60797/IRJ.2024.144.174.3>

Проблема для задачи многоклассовой классификации

Как упоминалось выше, функция softmax широко используется в задачах многоклассовой классификации. Из-за природы функции softmax все выходные значения будут представлять собой вероятности каждого класса. Независимо

от того, какой класс имеет наибольшее значение вероятности, система предсказывает, что входное изображение принадлежит именно этому классу.

Однако, при рассмотрении конкретного примера, такого как классификация знаковых зданий города Москвы [6], в которой набор данных включает 13 выходных классов, соответствующих 13 зданиям в городе, возникает проблема интерпретации результатов. Система классификации может точно предсказать здание, если входные данные изображения принадлежат одному из классов набора данных. Однако в случае если входные данные не соответствуют ни одному из классов набора данных (например, изображения реки, горы, человека, странного здания и т. д.), после прохождения через полносвязный слой с применением функции softmax, система все равно выдаст наибольшую вероятность определенного класса из набора, что может привести к ложным выводам и неправильной классификации. Данное явление требует системного контроля и обработки для предотвращения неоправданных результатов.

Сверимся с изображением храма Спаса на Крови, расположенным в городе Санкт-Петербург (см. рис. 4). Очевидно, что это изображение не соответствует ни одному классу в наборе данных (поскольку набор данных представляет собой здания в Москве).



Рисунок 4 - Образец изображения храма Спаса на Крови
DOI: <https://doi.org/10.60797/IRJ.2024.144.174.4>

Даже многие люди могут спутать храм Спаса на Крови с собором Василия Блаженного, потому что они имеют схожую архитектуру. Модель классификации, обученная с помощью трансферного обучения на основе модели InceptionV3 [7], также предсказала, что на изображении изображено здание Василия Блаженного. Ниже приведен результат модели, если последний слой использует функцию активации softmax:

```
[[4.2251e-07 8.3223e-06 1.1393e-05 1.3564e-05 2.9395e-06 0.00018633 0.00048449 1.2825e-05 2.5042e-08 4.021e-07 0.99911 0.0001692 2.4293e-09]]
```

index max: 10

label predict: St. Basils Cathedral

Из результатов видно, что модель, использующая функцию активации softmax, достигает очень высоких результатов (99,91%), поэтому нам очень сложно определить метки для изображений, которые не принадлежат ни к одному классу в наборе данных.

Решение

Чтобы решить эту проблему, мы применим функцию активации сигмовидной, а затем посмотрим на полученные результаты. Ниже приведены результаты модели, если к последнему слою применяется сигмовидная функция активации.

```
[[2.7557e-05 0.014681 9.0419e-05 0.0017351 0.0022757 0.00075092 0.00022904 6.3055e-06 0.00012885
2.1713e-05 0.31547 0.0082597 0.00029327]]
```

index max: 10

Таким образом, наибольшее значение в слоях классификации по-прежнему находится в 10-м слое (это Собор Василия Блаженного), но полученное значение значительно меньше (0,31547). Помните, что это число не является процентом вероятности для всех классификационных классов.

Поэтому мы можем установить пороговое значение, например 0,5. Если полученное максимальное значение больше 0,5, то мы можем предположить, что входное изображение принадлежит этому классу классификации, в противном случае мы можем заключить, что входное изображение не принадлежит ни одному классу классификации в наборе данных.

```
[[2.7557e-05 0.014681 9.0419e-05 0.0017351 0.0022757 0.00075092 0.00022904 6.3055e-06 0.00012885
2.1713e-05 0.31547 0.0082597 0.00029327]]
```

index max: 10

label predict: Unknown

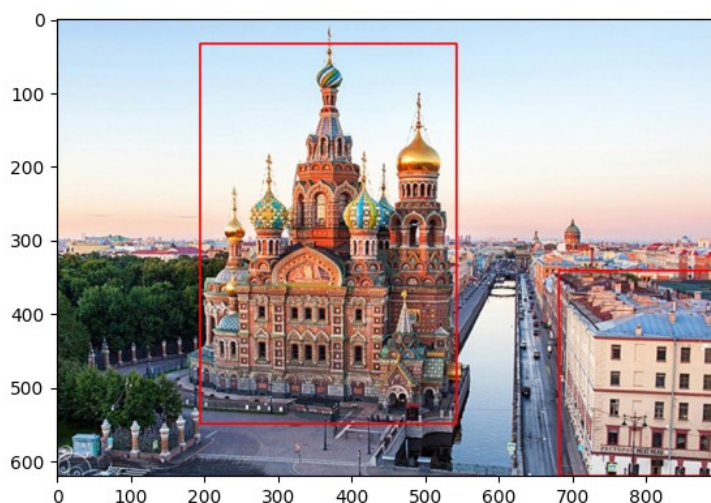


Рисунок 5 - Изображение с контуром рамки обнаружения здания
DOI: <https://doi.org/10.60797/IRJ.2024.144.174.5>

Заключение

В данной статье мы рассмотрели основы функций активации и их математические уравнения, сфокусировавшись на функциях softmax и сигмоиды. Функция активации играет важную роль в построении и обучении искусственных нейронных сетей, определяя их функциональные возможности. Мы изучили структуру и работу искусственного нейрона, а также принцип действия функций активации на примере сумматора.

Сигмоидальная функция широко используется в задачах бинарной классификации, преобразуя входное значение в диапазон от 0 до 1. Функция softmax, в свою очередь, применяется в многоклассовой классификации, преобразуя вектор значений в вероятностное распределение.

Однако, рассматривая конкретную проблему многоклассовой классификации зданий в городе Москве, мы обнаружили, что использование функции softmax может привести к ложным результатам, особенно когда входные данные не соответствуют ни одному классу в наборе данных. Это ставит под сомнение точность классификации и требует системного контроля и обработки таких случаев.

В заключении мы предложили решение данной проблемы, заключающееся в применении функции активации сигмоиды и установлении порогового значения для определения принадлежности входного изображения к классам классификации. Этот подход позволяет более точно оценивать результаты классификации и избегать ложных выводов.

Конфликт интересов

Не указан.

Рецензия

Кацко С.Ю., Сибирский государственный университет геосистем и технологий, Новосибирск, Российская Федерация
DOI: <https://doi.org/10.60797/IRJ.2024.144.174.6>

Conflict of Interest

None declared.

Review

Katsko S.Y., Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation
DOI: <https://doi.org/10.60797/IRJ.2024.144.174.6>

Список литературы / References

1. Красненко Д.М. Нейронные сети. Описание функций активации нейросети / Д.М. Красненко, А.С. Помогаев, А.С. Штыленко // Приоритетные направления инновационной деятельности в промышленности. Сборник научных статей IV международной научной конференции. — Казань, 2021. — С. 74-76.
2. Кобзев А.А. Анализ алгоритмов обучения нейронной сети / А.А., Кобзев, А.В. Лекарева, О.С. Сидорова // Современные наукоемкие технологии. — 2021. — № 6-1. — С. 23-28.
3. Ерш К.А. Сигмоидная функция активации как фильтр информации / К.А. Ерш // Традиционная и инновационная наука: история, современное состояние, перспективы. Сборник статей Международной научно-практической конференции. — 2019. — С. 29-31
4. Маршаков Д.В. Сравнение результатов нейросетевой классификации с применением SOFTMAX и функции расстояния / Д.В. Маршаков // Математические методы в технологиях и технике. — 2021. — № 8. — С. 75-78.
5. Костерин М.А. Нейросетевая классификация русскоязычных предложений по тональности на четыре класса / М.А. Костерин, И.В. Парамонов // Моделирование и анализ информационных систем. — 2022. — Т. 29. — № 2. — С. 116-133.
6. Building Moscow Image Dataset.. — 2024 — URL: <https://universe.roboflow.com/thanhngq/building-moscow> (accessed: 24.03.2024)
7. Lin C. Transfer Learning Based Traffic Sign Recognition Using Inception-v3 Model / C. Lin, L. Li, W. Luo [et al.] // Periodica Polytechnica Transportation Engineering. — DOI: 10.3311/pptr.11480
8. Иванова К.М. История храма Спаса на Крови. Гуманитарные науки в современном вузе: вчера, сегодня, завтра / К.М. Иванова, Д.В. Логинова // Материалы V международной научной конференции. — Санкт-Петербург, 2022. — Т. 2. — С. 901-904.
9. Колобаев М.Е. Современные архитектуры сверточных нейронных сетей, применяемые для задач классификации изображений / М.Е. Колобаев, Д.А. Прытыка, Д.О. Михайлова [и др.] // Мягкие измерения и вычисления. — 2023. — Т. 69. — № 8. — С. 69-78.
10. Хрипунов Д.Д. Сверточные нейронные сети как инструмент решения задачи классификации изображений / Д.Д. Хрипунов // Ломоносовские научные чтения студентов, аспирантов и молодых ученых - 2022. — Северный (Арктический) федеральный университет имени М.В. Ломоносова, 2022. — С. 597-600.

Список литературы на английском языке / References in English

1. Krasnenko D.M. Nejrornyie seti. Opisanie funkciij aktivacii nejroseti [Neural networks. Description of neural network activation functions] / D.M. Krasnenko, A.S. Pomogaev, A.S. SHtylenko // Prioritetnyye napravleniya innovacionnoj deyatel'nosti v promyshlennosti. Sbornik nauchnyh statej IV mezhdunarodnoj nauchnoj konferencii [In collection: Priority directions of innovation activity in industry. Collection of scientific articles of the IV International Scientific Conference]. — Kazan, 2021. — P. 74-76. [in Russian]
2. Kobzev A.A. Analiz algoritmov obucheniya nejronnoj seti [Analysis of neural network training algorithms] / A.A., Kobzev, A.V. Lekareva, O.S. Sidorova // Sovremennye naukoemkie tekhnologii [Modern science-intensive technologies]. — 2021. — № 6-1. — P. 23-28. [in Russian]

3. Ersh K.A. Sigmoidnaya funkciya aktivacii kak fil'tr informacii [Sigmoidal activation function as an information filter] / K.A. Ersh // Tradicionnaya i innovacionnaya nauka: istoriya, sovremennoe sostoyanie, perspektivy. Sbornik statej Mezhdunarodnoj nauchno-prakticheskoy konferencii [Traditional And Innovation Science: History, Modern State, Perspectives. Collection of articles of the International Scientific and Practical Conference]. — 2019. — P. 29-31 [in Russian]
4. Marshakov D.V. Sravnenie rezul'tatov nejrosetevoj klassifikacii s primeneniem SOFTMAX i funkcii rasstoyaniya [Comparison of neural network classification results using SOFTMAX and distance function] / D.V. Marshakov // Matematicheskie metody v tekhnologiyah i tekhnike [Mathematical methods in technologies and engineering]. — 2021. — № 8. — P. 75-78. [in Russian]
5. Kosterin M.A. Nejrosetevaya klassifikaciya russkoyazychnyh predlozhenij po tonal'nosti na chetyre klassa [Neural network classification of Russian-language sentences by tone into four classes] / M.A. Kosterin, I.V. Paramonov // Modelirovanie i analiz informacionnyh sistem [Modelling and analysis of information systems]. — 2022. — Vol. 29. — № 2. — P. 116-133. [in Russian]
6. Building Moscow Image Dataset.. — 2024 — URL: <https://universe.roboflow.com/thanhngq/building-moscow> (accessed: 24.03.2024)
7. Lin C. Transfer Learning Based Traffic Sign Recognition Using Inception-v3 Model / C. Lin, L. Li, W. Luo [et al.] // Periodica Polytechnica Transportation Engineering. — DOI: 10.3311/pptr.11480
8. Ivanova K.M. Istoriya hrama Spasa na Krovi. Gumanitarnye nauki v sovremennom vuze: vchera, segodnya, zavtra [History of the Church of the Saviour on the Blood. Humanities in the modern university: yesterday, today, tomorrow] / K.M. Ivanova, D.V. Loginova // Materialy V mezhdunarodnoj nauchnoj konferencii [Materials of the V international scientific conference]. — St. Petersburg, 2022. — Vol. 2. — P. 901-904. [in Russian]
9. Kolobaev M.E. Sovremennye arhitektury svertochnyh nejronnyh setej, primenyaemye dlya zadach klassifikacii izobrazhenij [Modern architectures of convolutional neural networks used for image classification tasks] / M.E. Kolobaev, D.A. Prytyka, D.O. Mihajlova [et al.] // Myagkie izmereniya i vychisleniya [Soft Measurements and Computation]. — 2023. — Vol. 69. — № 8. — P. 69-78. [in Russian]
10. Khripunov D.D. Svertochnye nejronnye seti kak instrument resheniya zadachi klassifikacii izobrazhenij [Converged neural networks as a tool for solving the problem of image classification] / D.D. Khripunov // Lomonosovskie nauchnye chteniya studentov, aspirantov i molodyh uchenyh - 2022 [Lomonosov Scientific Readings of Students, Postgraduates and Young Scientists - 2022]. — Lomonosov Northern (Arctic) Federal University, 2022. — P. 597-600. [in Russian]