

Трехмерная визуализация макропараметров обучения нейронных сетей

Ханиев Радмир Эмильевич

Кафедра проблем управления, Москва

E-mail: khaniev.raddark17@mail.ru

Соавторы: Дианэ С.А.К.

Обучение многослойных нейронных сетей прямого распространения (МНСПР) в задачах регрессии нелинейных функций или классификации данных является сложной многопараметрической, а зачастую и многокритериальной задачей. При этом можно выделить два уровня настройки МНСПР.

На микроуровне происходит автоматический поиск весовых коэффициентов нейросетевой модели с применением метода градиентного спуска [1]. На макроуровне реализуется выбор параметров структуры нейронной сети и алгоритма обучения, таких как общее количество нейронов, число используемых обучающих примеров, число эпох обучения. Для выбора оптимальных макропараметров обучения нейронной сети зачастую применяются интуитивные соображения эксперта и эмпирические расчетные формулы [2]. В этой связи целесообразно сделать процесс настройки более наглядным и удобным в отладке [3, 4]. Основными критериями, выбираемыми для оценки эффективности МНСПР, являются установившаяся ошибка E регрессии или классификации на тестовом множестве примеров, а также мера C вычислительной сложности нейронной сети. В настоящем докладе предлагается визуализировать ошибку обучения нейронной сети для многомерного пространства, состоящего из трех макропараметров – количества нейронов в скрытом слое, числа обучающих примеров и количества эпох обучения.

Основными требованиями, предъявляемыми к алгоритмическому обеспечению, являются экспериментальное получение оценок эффективности МНСПР с последующей записью их в файл, а также визуализация данных оценок для различных сочетаний входных макропараметров.

Структура ПО включает: процедуру обучения нейронной сети, нейронную сеть, таблицу результатов экспериментов, диапазоны вариации параметров, процедуру построения 1D-графика, процедуру построения 3D-графика, Matplotlib, PyGame, OpenGL, графические окна.

При реализации ПО использовался язык программирования Python [5]. Для визуализации результатов применялись графические библиотеки Matplotlib и OpenGL [6]. Построение трехмерной зависимости. Каждая точка, расположенная в участке дискретизированного трехмерного пространства, несет информацию о величине ошибки, полученную при обучении нейронной сети с макропараметрами, соответствующими координатам данной точки. Отметим, что для наглядности куб ошибок выводится на экран в анимированной форме, вращаясь вокруг вертикальной оси на 1° каждые 10 мс, что достигается изменением угла поворота α_y . В целях автоматизации выбора наилучших параметров экспертная функция оценки может быть сформирована в виде взвешенной суммы следующих параметров: функционал качества, количество эпох обучения, количество нейронов, количество эталонных точек.

Максимизация данного функционала для выбранного набора коэффициентов позволяет определить наилучшее сочетание макропараметров алгоритма обучения.

Трехмерный подход к отображению графика ошибок является достаточно эффективным средством оценки эффективности и анализа ошибок, возникающих при настройке нейронных сетей. В рамках исследования влияния нескольких макропараметров на качество обучения сети изучены частные одномерные зависимости ошибок RMSE и обобщенная трехмерная функция ошибок, которая учитывает количество нейронов в скрытом

слое, число обучающих примеров и количество эпох обучения. Анализ полученного трехмерного графика позволил сделать вывод о нелинейном характере многомерной зависимости и определить наилучшее сочетание макропараметров с использованием экспертной функции оценки.

Источники и литература

- 1 Хайкин С. Нейронные сети: полный курс. / 2-е изд. М.: Издательский дом "Вильямс", 2019. - 1104 с.
- 2 Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы // Пер. с польск. И.Д. Рудинского. – М.: Горячая линия – Телеком, 2006. – 452 с
- 3 Neural network assisted multiscale analysis for the elastic properties prediction of 3D braided composites under uncertainty [Электронный ресурс] / G. Balokas, S. Czichon, R. Rolfes // Composite Structures.2017.037. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0263822317302416> (дата обращения 15.06.2023)
- 4 Improving 3D convolutional neural network comprehensibility via interactive visualization of relevance maps: evaluation in Alzheimer's disease [Электронный ресурс] / M. Dyrba [и др.] // Alzheimer's Research & Therapy.2021.191. URL: <https://alzres.biomedcentral.com/articles/10.1186/s13195-021-00924-2> (дата обращения 17.06.2023)
- 5 Kuhlman D. A Python Book: Beginning Python, Advanced Python, and Python Exercises. – Boston: MIT, 2015. – 278 p.
- 6 Shreiner D. OpenGL programming guide: the official guide to learning OpenGL, versions 3.0 and 3.1. 7 изд. Addison Wesley, 2009. 1015 с.
- 7 Hazzat S., Saaidi A., Satori K. Multi-view passive 3D reconstruction: Comparison and evaluation of three techniques and a new method for 3D object reconstruction. Proc. of International Conference on Next Generation Networks and Services, NGNS. 2014. pp. 194-201.

Иллюстрации

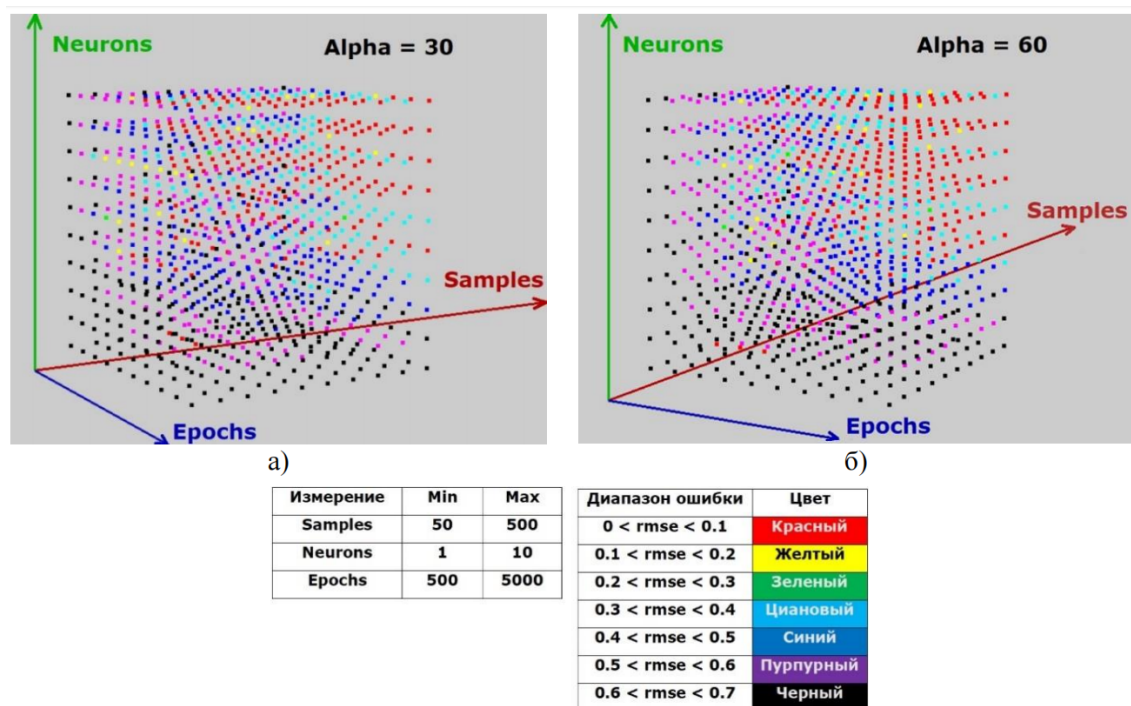


Рис. : 1. Трехмерная зависимость ошибки нейронной сети от макропараметров алгоритма обучения: а) визуализация при угле поворота $\alpha\gamma=30^\circ$; б) визуализация при угле поворота $\alpha\gamma=60^\circ$; в) легенда трехмерного графика