

РАЗРАБОТКА НЕЙРОСЕТЕВОЙ БИБЛИОТЕКИ ДЛЯ НЕЙРОННЫХ СЕТЕЙ ПРЯМОГО РАСПРОСТРАНЕНИЯ

Иванова Ю.А.¹, Борн Р.С.², Гордеев К.Е.³

¹ *Томский политехнический университет, ИШИТР, доцент, e-mail: jbolotova@tpu.ru*

² *Томский политехнический университет, ИШИТР, студент гр. 8В12, e-mail: rsb8@tpu.ru*

³ *Томский политехнический университет, ИШИТР, студент гр. 8В14, e-mail: keg6@tpu.ru*

Введение

Библиотека нейронных сетей – это программа, предназначенная для упрощения создания, обучения и развёртывания искусственных нейронных сетей. Эти библиотеки предоставляют инструменты и готовые компоненты, помогающие разработчикам быстро и эффективно создавать искусственные нейронные сети (ИНС). ИНС является моделью машинного обучения, частично воспроизводящей структуру и функции человеческого мозга. ИНС используются для извлечения закономерностей из данных, составления прогнозов и выполнения таких задач, как распознавание изображений и речи.

Одним из наиболее распространённых применений нейронных сетей является решение задачи классификации. Задача классификации предполагает контролируемое обучение, что означает, что алгоритм обучается на размеченных данных, содержащих правильные ответы. Цель задачи классификации состоит в отнесении объекта, представленного набором некоторых его признаков, к одной из заранее определённых категорий или классов.

Задача классификации нейронных сетей является важной частью области машинного обучения. Она позволяет нам распределять данные по различным классам на основе их особенностей, что имеет широкий спектр применения в различных областях. Используя библиотеки нейронных сетей и выполняя правильные шаги для обучения, мы можем создавать мощные и точные модели для задач классификации.

Целью нашей работы является разработка собственной библиотеки нейронных сетей прямого распространения. На основе разработанной библиотеки нами создана ИНС, способная классифицировать рукописные цифры и сигналы светофора с точностью 97.14% и 99.87% соответственно.

Описание алгоритма

Обучение нейронной сети для решения задачи классификации включает в себя несколько этапов. Первый шаг – сбор и предварительная обработка данных. Он включает в себя очистку данных, удаление любых выбросов и разделение данных на обучающий и тестовый наборы. Обучающий набор используется для обучения нейронной сети распознаванию паттернов, а тестовый набор – для оценки эффективности обученной модели.

Для подготовки данных обучающей и тестовой выборок первой нейронной сети для распознавания рукописных цифр был выбран открытый датасет MNIST [3], содержащий изображения рукописных цифр размером 28×28 пикселей. Общее количество изображений – 70000, из которых 60000 составляют обучающую выборку и 10000 – тестовую выборку. Примеры изображений представлены на Рис. 1.

Для подготовки данных обучающей и тестовой выборок нейронной сети, предназначенной для распознавания сигналов светофора, был выбран открытый датасет WPI Traffic Light Dataset [4], содержащий изображения красных и зелёных сигналов светофора. Изображения были приведены к единому размеру 32×32 пикселя. Общее количество изображений – 7818, из которых 6254 составляют обучающую выборку и 1564 – тестовую. Примеры изображений представлены на Рис. 2.

Следующий шаг – разработка архитектуры нейронной сети. Он включает в себя выбор количества слоёв, количества нейронов в каждом слое и функций активации. Архитектура нейронной сети может оказать значительное влияние на её производительность, поэтому важно выбрать наиболее подходящую структуру для решения поставленной задачи. Для обеих нейронных сетей экспериментально была выбрана двухслойная архитектура. Для первой нейроны распределены как 784×300×10, для второй – 3072×1500×2.

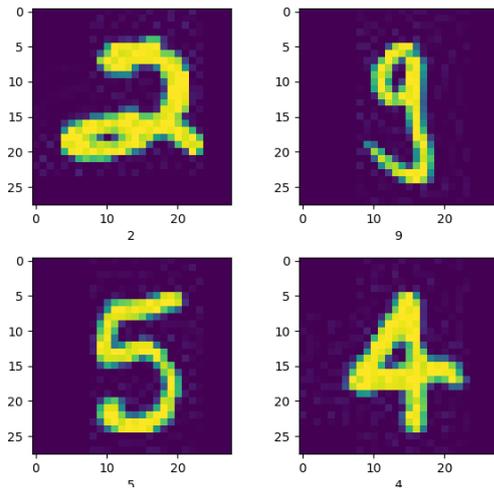


Рис. 1. Примеры датасета MNIST

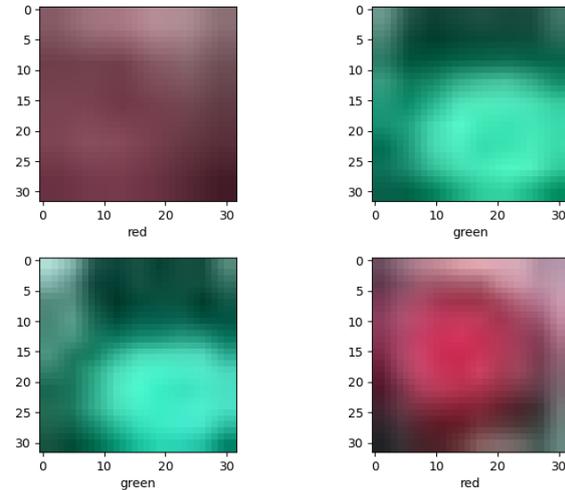


Рис. 2. Примеры датасета WPI Traffic Light Dataset

Обучение нейронной сети осуществляется с помощью алгоритма обратного распространения ошибки. Он включает в себя регулировку весов и смещений нейронов в сети с целью минимизации среднеквадратичной ошибки (MSE) между прогнозируемым и фактическим выходом. Этот процесс повторяется до тех пор, пока сеть не сможет с заданной точностью предсказать выходные классы для тестового набора, либо пока не закончится отведённое на обучение время.

Работа ИНС при разных параметрах представлена в Таблице 1. Исходя из полученных данных, первоначально выбранная скорость обучения, равная 0.8, даёт низкое качество обучения, из чего следует, что для данной нейронной сети скорость обучения необходимо снизить.

Таблица 1

Показатели работы ИНС при различных параметрах

Кол-во эпох	Скорость обучения	Точность ИНС для распознавания рукописных цифр, %	Точность ИНС для распознавания сигнала светофора, %
1	0,05	91,72	99,68
3	0,8	88,03	50
3	0,05	95,01	99,48
5	0,15	96,69	99,87
7	0,8	89,50	50
10	0,15	97,14	99,87

Графики функции потерь MSE в зависимости от количества эпох, полученные в ходе обучения, представлены на Рис. 3, 4. По оси абсцисс откладывается номер эпохи, по оси ординат – ошибка сети. На данных графиках представлено снижение потерь моделей при увеличении количества эпох.

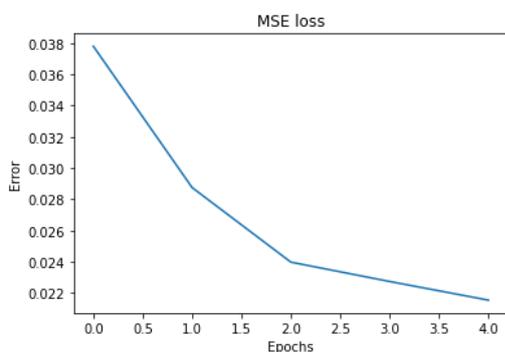


Рис. 3. График функции потерь для ИНС распознающей сигналы светофора

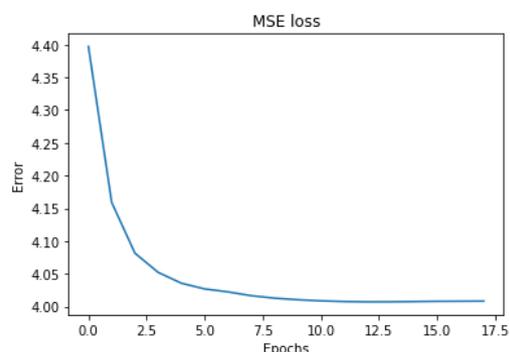


Рис. 4. График функции потерь для ИНС распознающей рукописные цифры

Заключение

В результате научно-исследовательской работы была разработана нейросетевая библиотека. На основе неё были разработаны 2 нейронных сети прямого распространения, решающие задачу распознавания рукописных цифр (MNIST) с точностью 97.17% и средней ошибкой 4%, что оказывается лучшим результатом, чем точность, достигнутая сетями с подобной архитектурой [2], и задачу распознавания сигнала светофора с точностью 99.87%.

Список использованных источников

1. В. Г. Спицын, Ю.Р. Цой. Интеллектуальные системы. – Томск: ТПУ, 2012. – 176 с.
2. THE MNIST DATABASE of handwritten digits. URL: <http://yann.lecun.com/exdb/mnist/> (дата обращения: 09.02.2023).
3. WPI Traffic Light Dataset. URL: <https://computing.wpi.edu/dataset.html> (дата обращения: 09.02.2023).
4. Метрики классификации и регрессии. URL: <https://academy.yandex.ru/handbook/ml/article/metriki-klassifikacii-i-regressii> (дата обращения: 09.02.2023).