МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования   
**«Национальный исследовательский   
Нижегородский государственный университет им. Н.И. Лобачевского»**

**(ННГУ)**

**Институт информационных технологий, математики и механики**

Направление подготовки: «Прикладная математика и информатика»

Магистерская программа: «Математическая кибернетика»

Образовательный курс «Методы глубокого обучения для решения задач компьютерного зрения»

**ОТЧЕТ**

по лабораторной работе №1

**Реализация метода обратного распространения ошибки для двуслойной полностью связанной нейронной сети**

**Выполнила:**

студентка группы 381603м1

Шестакова Светлана

Нижний Новгород

2018

Содержание

[1 Постановка задачи 3](#_Toc506460684)

[1.1 Цель работы 3](#_Toc506460685)

[1.2 Задачи работы 3](#_Toc506460686)

[2 Описание метода обратного распространения ошибки. Вывод математических формул. 4](#_Toc506460687)

[2.1 Математическое объяснение метода. Постановка задачи оптимизации 4](#_Toc506460688)

[2.2 Обратное распространение 5](#_Toc506460689)

[3 Алгоритм метода обратного распространения ошибки. 7](#_Toc506460690)

[4 Описание программной реализации. 8](#_Toc506460691)

[4.1 Структура проекта 8](#_Toc506460692)

[4.2 Руководство пользователя 8](#_Toc506460693)

[5 Результаты 9](#_Toc506460694)

# Постановка задачи

## Цель работы

## Целью лабораторной работы является изучение и реализация метода обратного распространения ошибки для обучения глубоких нейронных сетей на примере двухслойной полностью связанной сети (один скрытый слой), используянабор данных MNIST.

## Задачи работы

В соответствии с поставленной целью,необходимо решение следующих задач:

1. Изучить общую схему метода обратного распространения ошибки.
2. Вывести необходимые математические формулы для вычисления градиентов функции ошибки по параметрам нейронной сети и формулу коррекции весов.
3. Спроектировать и разработать программную реализацию метода, который позволяет работать с набором данных MNIST
4. Произвести тестирование разработанной программной реализации.
5. Нахождение оптимальных параметров функционирования.

# Описание метода обратного распространения ошибки. Вывод математических формул.

## Математическое объяснение метода. Постановка задачи оптимизации

Введем следующие буквенные обозначения:

* 𝑁 – количество входных нейронов;
* 𝑀 – количество выходных нейронов;
* 𝐾 – количество нейронов на скрытом слое;
* 𝐿 – количество обучающих примеров.

В качестве функции ошибки целесообразно рассмотреть кросс-энтропию:

, ,

где

– множество обучающих примеров,

– выход нейронной сети, полученный для входного примера.

.

Предположим, что режим обучения является последовательным. Тогда корректировка весов должна выполняться после прохода каждого примера обучающей выборки.

Рассмотрим пример:

,

,

.

В этом случае, а функция ошибки принимает следующий вид:

.

Введем следующие обозначения веса синаптических связей:

– от входных нейронов к нейронам скрытого слоя,

– от нейронов скрытого слоя к выходным нейронам нашей сети. Выходной сигнал нейрона скрытого слоя вычисляется следующим образом:

, где 𝜑 – функция активации на скрытом слое,

- взвешенная сумма входных сигналов.

Сигнал выходного нейрона можно определить как

, где ℎ - функция активации на последнем слое,

– взвешенная сумма сигналов со скрытого слоя.

В качестве функции активации на выходном слое рассмотрим функцию*softmax*:

Таким образом,

,

.

Глядя на полученную функцию ошибки, можно сказать, что задача обучения нейронной сети сводится к задаче оптимизации функции ошибки по всем весам сети

𝐸(𝑤) → 𝑚𝑖𝑛𝑤.

## Обратное распространение

Метод обратного распространения ошибки определяет способ проведения изменения параметров сети 𝑤.

Для этого можно использовать градиентные методы оптимизации. Производная целевой функции по параметрам последнего слоя вычисляется по следующей формуле:

,

,

В рассматриваемой задаче:

.

Таким образом:

.

Производная целевой функции по параметрам скрытого слоя вычисляется по формуле:

Таким образом:

.

В случае, если на скрытом слое функция активации являетсягиперболическим тангенсом: 𝜑(𝑓𝑠 ) = 𝑡ℎ(𝑓𝑠), то

Градиент может быть выражен следующим образом:

,

.

Согласно градиентным методам на каждом шаге 𝑟 + 1 обучения сети необходимопроизводить коррекцию весов следующим образом:

,

,

где𝜂 – скоростьобучения.

# Алгоритм метода обратного распространения ошибки.

Алгоритм метода обратного распространения ошибки фактически состоит из шести шагов:

1. Инициализация весов 𝑤 некоторыми значениями
2. Прямой проход нейронной сети
3. Обратный проход
4. Шаги 3-5 повторяются до тех пока, пока не выполнится критерий остановки. Как правило, это либо максимальное число эпох либо достигнутая точность обучения.

Прямой проход.

На вход подается𝑥𝑖. Необходимо вычислить значения выходных сигналов нейронов скрытого слоя - количество нейронов на скрытом слое и значение производной функции активации на скрытом слое .

Вычислить выходные сигналы нейронов последнего слоя

– количество классов изображений.

Коротко, его можно изобразить как: 𝑥𝑖 → 𝑣𝑠 ,

Обратный проход:

Вычислим значения градиентов целевой функции, начиная с конца:

*for*

,

Скрытый слой:

*for*

По дугам:

,

,

# Описание программной реализации.

## Структура проекта

Проект был разработан в MSVisualStudio2015 и содержит следующие файлы:

* ReaderFile.h – содержит функции для работы с набором данных MNIST
* NeuralNet.h – содержит описание класса нейронной сети
* NeuralNetwork.cpp –содержит методы для работы с нейронной сетью
* Main.cpp – главный файл проекта

## Руководство пользователя

1. Path to MNIST train-images – обязательный параметр
2. Path to MNIST train-labels – обязательный параметр
3. Path to MNIST test-images – обязательный параметр
4. Path to MNIST train-labels – обязательный параметр
5. numberEpochs – число эпох для расчета (по умолчанию = 15)
6. crossError – точность обучения для критерия остановки (по умолчанию = 0.005)
7. learnRate – скорость обучения (по умолчанию = 0.01)
8. numberHiddenNeurons – число нейронов скрытого слоя (по умолчанию = 200)

# Результаты

Было разработано приложение, позволяющее обучать и тестировать двухслойную нейронную сеть с использование набора данных MNIST.

Требуемое значение кросс-энтропии для всех экспериментов было указано 0.005, скорость обучения 0.01. Результаты представлены в таблице ниже.

|  |  |  |  |
| --- | --- | --- | --- |
| Количество эпох | Количество нейронов | Точность на тестовой выборке | Точность на обучающей выборке |
| 10 | 100 | 0.9541 | 0.9599 |
| 15 | 100 | 0.953 | 0.9644 |
| 20 | 200 | 0,9518 | 0.9657 |