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**

**Кафедра математического обеспечения и суперкомпьютерных технологий**

Направление подготовки: «Фундаментальная информатика

и информационные технологии»

Магистерская программа: «Компьютерная графика»

**Отчет по лабораторной работе**

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

Выполнила:

студентка группы 381606м2

Шерстнева А.Д.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Нижний Новгород

2017

Оглавление

[Постановка задачи 3](#_Toc501530825)

[Метод обратного распространения ошибки 4](#_Toc501530826)

[Программная реализация 7](#_Toc501530827)

[Результаты экспериментов 7](#_Toc501530828)

# Постановка задачи

* Изучить метод обратного распространения ошибки;
* Вывести математические формулы для вычисления градиентов функции ошибки по параметрам нейронной сети и формул коррекции весов;
* Разработать программную реализацию, провести эксперименты.
* Подготовить отчет по проделанной работе.

# Метод обратного распространения ошибки

**Описание:**

Идея метода состоит в распространении сигналов ошибки от выходов к входам сети. У нас есть множество входов . Обозначим вес i-го и j-го ребра через , выход j-го нейрона обозначим через . Так же имеем ожидаемый результат – вектор . Чтобы узнать, как текущий результат отличается от ожидаемого, будем использовать функцию ошибки кросс-энтропию:

Метод обратного распространения ошибки состоит из 2х частей. Перед началом обучения необходимо инициализировать веса и смещения малыми значениями. Далее начинается обучение сети.

**Прямой проход:**

1. Подаем на вход вектор входных сигналов .
2. Каждый нейрон входного слоя отправляет полученный сигнал всем нейронам скрытого слоя.
3. Каждый скрытый нейрон суммирует взвешенные входящие сигналы

, а затем применяет активационную функцию:

1. Каждый выходной нейрон суммирует взвешенные входящие сигналы

, а затем применяет активационную функцию:

**Обратный проход:**

1. Каждый выходной нейрон получает целевые значения и вычисляет ошибку:

Используем получившиеся значения:

1. Каждый скрытый нейрон суммирует входящие ошибки (от нейронов в последующем слое) и вычисляет величину ошибки, умножая полученное значение на производную активационной функции:

где – поправка, вычисленная для к-го узла.

Таким образом получаем следующую ошибку для нейронов скрытого слоя:

И ошибку для нейронов выходного слоя равную

1. Каждый выходной и скрытый нейрон будет изменять веса своих скрытых связей по формуле:

где , a – скорость обучения нейронной сети, – ошибка, – выход предыдущего слоя

# Программная реализация

Программная реализация нейронной сети содержит следующие файлы:

* NeuronNetw.h – описание класса нейронной сети
* NeuronNetw.cpp – реализация функционала для работы с сетью.
* Main.cpp – приложение для запуска сети. В нем происходит загрузка данных MNIST, создание нейронной сети и подача данных в созданную сеть для обучения.

# Результаты экспериментов

Ниже в таблице представлены результаты обучения и тестирования сети на наборе данных MNIST. В зависимости от зависимости параметра скорости обучения изменялась точность выдаваемых результатов.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Кол-во нейронов на скр. слое | Скорость обучения | Кол-во эпох | Кол-во данных | Мин. ошибка | Точность  (1 – min\_er) | Энтропия | Кол-во верно угаданных данных |
| Train Data | | | | | | | |
| 200 | 0,001 | 15 | 60 000 | 0,058 | 0,942 | 0,1983 | 56 517 |
| 200 | 0,005 | 15 | 60 000 | 0,022 | 0,9776 | 0,086 | 58 652 |
| 200 | 0,008 | 15 | 60 000 | 0,022 | 0,978 | 0,077 | 58 682 |
| Test Data | | | | | | | |
| 200 | 0,001 | 15 | 10 000 | 0,043 | 0,9571 | 0,157 | 9 571 |
| 200 | 0,005 | 15 | 10 000 | 0,011 | 0,9895 | 0,046 | 9 895 |
| 200 | 0,008 | 15 | 10 000 | 0,006 | 0,9944 | 0,028 | 9 944 |

Т.о. лучший результат сеть дала при скорости обучения равной 0,008:

97% - точность на тренировочной выборке

99% - точность на тестовой выборке