# Лабораторная работа №1

# Метод обратного распространения ошибки для искусственной нейронной сети

## Постановка задачи

Изучение метода обратного распространения ошибки для двухслойной классифицирующей нейронной сети , вывод формул метода, программная реализация и проведение тестов.

Для тестирования сети используется задача классификации изображений рукописных цифр: <http://yann.lecun.com/exdb/mnist/>

## Описание нейронной сети

Сеть двухслойная, во входном слое 28\*28 = 784 нейронов по числу пикселей входных изображений, в выходном - 10 по числу классов.

В скрытом слое используется логистическая функция активации

D:\Университет\Магистратура\Глубокое обучение\отчет\logistic.png

D:\Университет\Магистратура\Глубокое обучение\отчет\logistic der.png

В выходном слое - функция softmax

D:\Университет\Магистратура\Глубокое обучение\отчет\softmax.png

D:\Университет\Магистратура\Глубокое обучение\отчет\softmax der 1.png

D:\Университет\Магистратура\Глубокое обучение\отчет\softmax der 2.png

## Метод обратного распространения ошибки

### Обозначения

x - входной вектор сети

y - выходной вектор сети

N - количество нейронов входного слоя

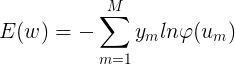
K - количество нейронов скрытого слоя

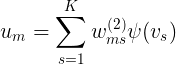
M - количество нейронов выходного слоя

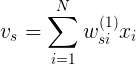
- веса скрытого слоя

- веса выходного слоя

E(w) - функция ошибки (кросс-энтропия)

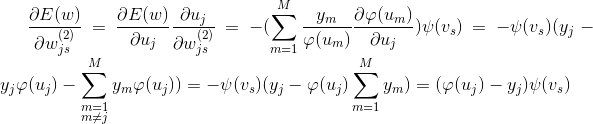




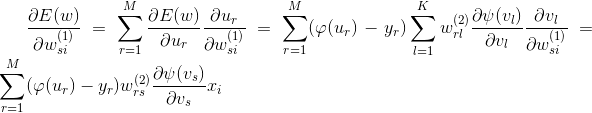


### Формулы градиентов функции ошибки

По весам выходного слоя:



По весам скрытого слоя:



### Общая схема алгоритма

1. пока не превышено заданное число итераций
   1. для очередной обучающей пары (x, y) вычислить выходы скрытого и выходного слоя, вычислить производную на скрытом слое
   2. вычислить градиенты функции ошибки по приведённым выше формулам
   3. изменить веса сети по формуле градиентного спуска, - параметр скорости обучения

## Программная реализация

Класс DigitsDataset предоставляет функции для работы с обучающими и тестовыми данными. Следующие функции предоставляют последовательный доступ к парам «изображение – метка класса» из обучающей и тестовой выборки соотвественно:

* get\_next\_train
* get\_next\_test
* get\_next\_train\_col
* get\_next\_test\_col

Первые две функции возвращают изображение в обычном виде – как двумерный массив. Вторые две возвращают изображение, все пиксели которого записаны в один столбец.

Функция reset позволяет вернуться к началу выборки.

Класс NeuralNet представляет реализацию нейронной сети и содержит следующие публичные функции:

* get\_output – возвращает выход сети для заданного входа
* train – запускает алгоритм обратного распростанения ошибки для обучения сети
* error\_test – возвращает ошибку в % (как соотношение неправильно классифицированных к общему числу примеров) сети на тестовой выборке
* error\_train - возвращает ошибку в % (как соотношение неправильно классифицированных к общему числу примеров) сети на обучающей выборке
* cross\_entropy – возвращает значение функции ошибки (кросс-энтропии)

## Результаты тестов

Ошибки в % на тестовой и обучающей выборках при разных числе итераций и параметре скорости обучения (49 нейронов в скрытом слое):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 500 | | 1000 | | 2000 | | 4000 | | 8000 | | 16000 | | 32000 | |
|  | train | test | train | test | train | test | train | test | train | test | train | test | train | test |
| 0.0125 | 60.10 | 59.62 | 48.70 | 48.55 | 34.69 | 34.94 | 24.64 | 23.89 | 19.51 | 18.24 | 14.24 | 13.81 | 10.28 | 9.59 |
| 0.025 | 51.30 | 50.85 | 42.53 | 43.08 | 25.48 | 25.05 | 18.85 | 18.04 | 14.55 | 13.72 | 12.46 | 12.38 | 8.94 | 9.03 |
| 0.05 | 40.48 | 40.17 | 31.72 | 31.49 | 20.71 | 20.36 | 15.26 | 14.6 | 13.38 | 13.07 | 10.48 | 10.32 | 7.90 | 7.89 |
| 0.1 | 30.33 | 30.77 | 28.08 | 27.79 | 20.01 | 19.36 | 14.09 | 13.69 | 13.40 | 12.97 | 9.36 | 9.68 | 7.11 | 7.12 |
| 0.2 | 35.03 | 34.82 | 28.63 | 28.53 | 20.13 | 19.38 | 14.70 | 14.09 | 14.51 | 14.83 | 11.38 | 11.57 | 9.05 | 9.06 |
| 0.4 | 43.26 | 44.39 | 41.87 | 43.2 | 19.53 | 19.74 | 26.05 | 25.31 | 20.83 | 19.88 | 16.58 | 16.78 | 14.84 | 14.42 |

Ошибки в % на тестовой и обучающей выборках при разных числе итераций и параметре скорости обучения (294 нейронов в скрытом слое):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 500 | | 1000 | | 2000 | | 4000 | | 8000 | | 16000 | | 32000 | |
|  | train | test | train | test | train | test | train | test | train | test | train | test | train | test |
| 0.0125 | 46.63 | 46.52 | 37.76 | 38.8 | 25.75 | 24.77 | 19.06 | 19.22 | 13.41 | 13.53 | 11.49 | 11.71 | 9.3 | 9.52 |
| 0.025 | 41.64 | 42.3 | 27.49 | 27.34 | 20.19 | 19.68 | 16.63 | 16.04 | 19.34 | 19.06 | 11.79 | 11.47 | 7.82 | 8.15 |
| 0.05 | 35.28 | 35.02 | 25.87 | 25.35 | 20.81 | 19.97 | 13.70 | 13.85 | 15.18 | 15.24 | 11.37 | 11.46 | 6.83 | 7.34 |
| 0.1 | 33.61 | 33.21 | 24.55 | 23.99 | 16.01 | 15.84 | 14.18 | 13.72 | 12.64 | 12.32 | 10.79 | 10.81 | 6.23 | 6.34 |
| 0.2 | 37.09 | 37.34 | 32.66 | 33.14 | 16.16 | 15.32 | 18.13 | 18.47 | 19.41 | 19.6 | 9.22 | 9.53 | 9.1 | 8.91 |
| 0.4 | 40.89 | 41.72 | 28.83 | 28.9 | 22.39 | 21.02 | 22.55 | 22.23 | 30.89 | 31.6 | 26.78 | 27.44 | 31.32 | 31.31 |

Наилучшие результаты из проведённых тестов:

* на тестовой: 6.34%
* на обучающей: 6.23%

были получены при следующих параметрах:

* нейронов в скрытом слое: 294
* скорость обучения: 0.1
* итераций: 32000

График тестовой ошибки от числа итераций:

График тестовой ошибки от параметра скорости: