Оглавление

[Постановка задачи 2](#_Toc500710504)

[Функция активации 2](#_Toc500710505)

[Кросс-энтропия 2](#_Toc500710506)

[Алгоритм работы программы 3](#_Toc500710507)

[Данные для обучения 4](#_Toc500710508)

[Расчет ошибки 4](#_Toc500710509)

[Результаты работы 4](#_Toc500710510)

[Заключение 5](#_Toc500710511)

# Постановка задачи

Необходимо написать программу, реализующую трехслойный перцептрон. Проверить работоспособность нейронной сети на задаче классификации рукописных цифр из набора данных MNIST. В качестве функции активации на втором слое используется функция softmax. В качестве функции ошибки используется кросс-энтропия.

# Функция активации

Когда задача классификации требует распознавания для более чем двух классов, в качестве функции активации обычно применяется softmax для результирующего слоя. Softmax функция предоставляет способ для предсказания распределения вероятности между классами. В качестве функции ошибки используется кросс-энтропия.

Softmax функция активации для i-того результата:

, где - выходные значения слоя. – число нейронов на слое.

Функция активации на скрытом слое – сигмоида:

# Кросс-энтропия

Кросс энтропия в качестве функции ошибки для распознавания на более чем два класса:

, где – желаемый результирующий вектор, – полученный в результате вычисления результат.

Вычисление градиента:

Градиент для весов на последнем слое:

Градиент для нейронов скрытого слоя:

# Алгоритм работы программы

1. Считывание данных для обучения
2. Задание начальных значений – число нейронов скрытого слоя (), скорость обучения (), генерируются малые случайные значения для весов ребер ().
3. Обучение. Для каждого изображения из обучающей выборки повторить:
   1. Подать значения интенсивностей на вход и подсчитать значения на последнем слое.
   2. Применить функцию активации к выходным значениям на последнем слое:
   3. Рассчитать градиент на последнем слое:
   4. Рассчитать градиент на скрытом слое:
   5. Пересчитать веса ребер:
4. Тестирование. Для каждого изображения из тестовой выборки повторить:
   1. Подать значения интенсивностей на вход и подсчитать значения на последнем слое.
   2. Выбрать класс, имеющий максимальное значение на выходном слое.
   3. Сравнить выбранный класс с правильным ответом. Если значения не совпадают, увеличить счетчик ошибок.
5. Расчет процента ошибки.

# Данные для обучения

Для обучения перцептрона используется набор из 60000 изображений рукописных цифр MNIST. Сеть обучается одним полным последовательным проходом по всем данным из обучающей выборки. При увеличении числа проходов по данным незначительно увеличивается точность.

# Расчет ошибки

Для проверки работоспособности нейронной сети используется набор из 10000 тестовых изображений MNIST. В качестве результата классификатора выбирается класс, имеющий наибольшее значение после применения функции softmax к выходу последнего слоя. Полученный результат сравнивается с истинным значением. При несоответствии результата увеличивается счетчик ошибок. Полученное после полного тестирования количество ошибок делится на общее количество тестовых изображений для получения доли ошибки. Формула расчета вероятности ошибки в процентах:

# Результаты работы

Для тестирования работы перцептрона необходимо установить параметры обучения: скорость обучения и число нейронов на скрытом слое . В данном случае скорость обучения = 0.1. Число нейронов скрытого слоя = 300. Данные параметры можно указать в качестве параметров запуска.

В результате работы нейронной сети из 10000 изображений 490 были классифицированы неправильно. Таким образом, шанс на ошибку составил 4.9%. Точность классификатора – 0.951.

Так же был проведен эксперимент, когда обучение проводилось тремя проходами по тестовым данным. В таком случае ошибка составила 4.04%, точность классификатора – 0.9596

# Заключение

Написанный трехслойный перцептрон способен успешно справиться с задачей классификации рукописных цифр из набора данных MNIST. Параметры сети можно настроить для увеличения точности определения. Также для увеличения точности классификации можно производить обучение в несколько проходов по тестовой выборке.