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: «Вычислительные методы и суперкомпьютерные технологии»

Образовательный курс «Глубокое обучение»

**ОТЧЕТ**

по лабораторной работе №2

**Разработка полностью связанной нейронной сети**

**Выполнили:**

студенты группы 381703-3м

Гладкова Татьяна

Крутоборежская Ирина

Крюкова Полина

Подчищаева Мария

Нижний Новгород

2018

Содержание

[Цели 3](#_Toc530326965)

[Задачи 4](#_Toc530326966)

[Решаемая задача 5](#_Toc530326967)

[Выбор библиотеки 6](#_Toc530326968)

[Метрика качества решения задачи 6](#_Toc530326969)

[Тренировочные и тестовые наборы данных 6](#_Toc530326970)

[Конфигурации нейронных сетей 7](#_Toc530326971)

[Описание содержимого директории src 8](#_Toc530326972)

[Результаты экспериментов 8](#_Toc530326973)

[Анализ результатов 9](#_Toc530326974)

[Выводы 10](#_Toc530326975)

[Литература 11](#_Toc530326976)

# Цели

1. ***Цель*** настоящей работы состоит в том, чтобы получить базовые навыки работы с одной из библиотек глубокого обучения (Caffe, Torch, TensorFlow, MXNet или какая-либо другая библиотека на выбор студента) на примере полностью связанных нейронных сетей.

# Задачи

1. Выполнение практической работы предполагает решение ***следующих задач***:
2. 1. Выбор библиотеки для выполнения практических работ курса.
3. 2. Установка выбранной библиотеки на кластере (параметры аутентификации и инструкция по работе с кластером выложена в отдельной задаче в системе redmine).
4. 3. Проверка корректности установки библиотеки. Разработка и запуск тестового примера сети, соответствующей логистической регрессии, для решения задачи классификации рукописных цифр набора данных MNIST (пример разобран в лекционных материалах).
5. 4. Выбор практической задачи компьютерного зрения для выполнения практических работ.
6. 5. Разработка программ/скриптов для подготовки тренировочных и тестовых данных в формате, который обрабатывается выбранной библиотекой.
7. 6. Разработка нескольких архитектур полностью связанных нейронных сетей (варьируются количество слоев и виды функций активации на каждом слое) в формате, который принимается выбранной библиотекой.
8. 7. Обучение разработанных глубоких моделей.
9. 8. Тестирование обученных глубоких моделей.
10. 9. Публикация разработанных программ/скриптов в репозитории на GitHub.
11. 10. Подготовка отчета, содержащего минимальный объем информации по каждому этапу выполнения работы.

# Решаемая задача

Была выбрана задача бинарной классификации: «кошки» - «собаки». Были использованы картинки из наборов данных https://www.kaggle.com/tongpython/cat-and-dog и [https://www.kaggle.com/c/dogs-vs-cats/data](https://vk.com/away.php?to=https%3A%2F%2Fwww.kaggle.com%2Fc%2Fdogs-vs-cats%2Fdata&cc_key=). Получившийся набор состоит из

35029 изображений.

|  |  |
| --- | --- |
|  |  |
| рис. 1 Пример изображения из класса «кошки» | рис. 2 Пример изображения из класса «собаки» |

С помощью скрипта на python данные были преобразованы к размеру 128×128. С помощью скрипта im2rec.py, который входит в библиотеку MXNet, изображения были сконвертированы в формат .rec.

# Выбор библиотеки

Для выполнения лабораторных работ была выбрана библиотека MXNet для языка программирования Python.

На этапе проверки корректности установки библиотеки была выполнена разработка и запуск тестового примера сети для решения задачи классификации рукописных цифр набора данных MNIST. Достигнута точность 0.9225.

# Метрика качества решения задачи

В качестве метрики точности решения используется отношение угаданных животных ко всем в тестовой выборке, т.е :

# Тренировочные и тестовые наборы данных

В качестве тренировочной выборки используем тренировочную выборку первого и второго наборов данных, всего 16500 изображений котов и 16505 изображений собак. В качестве тестовой выборки используем тестовую выборку только из первого набора данных, т.к. во втором наборе данных тестовая выборка не разбита на изображения котов и собак. Всего в тестовой выборку 2042 изображения, котов и собак поровну.

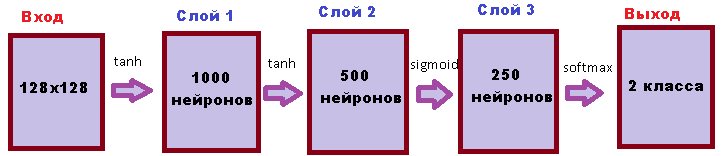
# Конфигурации нейронных сетей

В данной работе были рассмотрены четыре конфигурации полносвязных нейронных сетей с 3-мя и 4-мя скрытыми слоями.

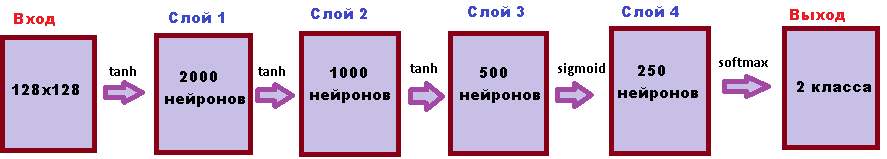
Активационная функция на слоях выбирается из следующих:

На выходном слое:

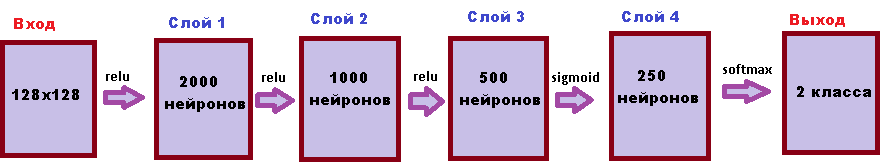
1.



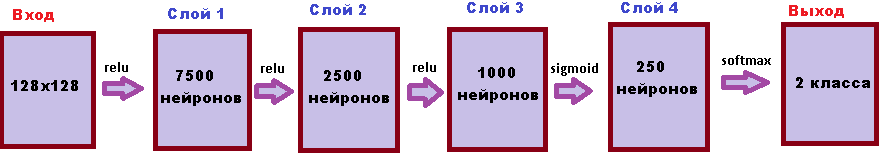
2.



3.



4.



# Описание содержимого директории src

В директории расположены четыре конфигурации построенных полносвязных нейронных сетей. Соответствия построенных конфигураций и конфигураций в директории:

* fcnn\_tts.py – первая нейронная сеть tanh-tanh-sigmoid, 1000-500-250
* fcnn\_ttts.py – вторая нейронная сеть tanh-tanh-tanh-sigmoid, 1000-500-250
* fcnn\_rrrs1.py – третья нейронная сеть relu-relu-relu-sigmoid, 2000-1000-500-250
* fcnn\_rrrs2.py – четвертая нейронная сеть relu-relu-relu-sigmoid, 7500-2500-1000-250

# Результаты экспериментов

В работе были рассмотрены 4 конфигурации.

Общие данные для всех экспериментов:

количество эпох = 10,

скорость обучения = 0.001.

Time – время обучения.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **№** | **Количество скрытых слоев** | **Количество нейронов на скрытых слоях** | **Функции активации** | **Результат** | | |
| Train-accuracy | Validation-accuracy | Time, sec |
| 1 | 3 | 1000-500-250 | tanh-tanh-sigmoid | 0.56 | 0.5 | 48.86 |
| 2 | 4 | 2000-1000-500-250 | tanh-tanh-tanh-sigmoid | 0.57 | 0.5 | 49.62 |
| 3 | 4 | 2000-1000-500-250 | relu-relu-relu-sigmoid | 0.79 | 0.74 | 47.29 |
| 4 | 4 | 7500-2500-1000-250 | relu-relu-relu-sigmoid | 0.83 | 0.78 | 165.79 |

# Анализ результатов

Нейронные сети с функцией активации ReLU показывают результат лучше, чем нейронные сети с функцией сигмоидальной функцией активации.

Рассмотрим функцию активации, которая представляется суммой нескольких логистических сигмоидов:

Построенную сумму можно приблизить интегралом:

Приведенный ряд сигмоидальных функций более выразителен и может быть приближен .

Указанный логарифм похож на ReLU.

Преимущества ReLU:

1. Вычисление сигмоиды и гиперболического тангенса требует ресурсоёмких операций, таких как возведение в степень, в то время как ReLU не подвержен насыщению.
2. Применение ReLU существенно повышает скорость стохастического градиентного спуска по сравнению с сигмоидой и гиперболическим тангенсом. Это обусловлено линейным характером и отсутствием насыщения данной функции.

Недостатки ReLU:

К сожалению, ReLU не всегда достаточно надёжны и в процессе обучения могут выходить из строя. Например, большой градиент, проходящий через ReLU, может привести к такому обновлению весов, что данный нейрон никогда больше не активируется. Если это произойдет, то, начиная с данного момента, градиент, проходящий через этот нейрон, всегда будет равен нулю. Соответственно, данный нейрон будет необратимо выведен из строя. Например, при слишком большой скорости обучения, может оказаться, что до 40% ReLU никогда не активируются. Эта проблема решается посредством выбора надлежащей скорости обучения.

# Выводы

Наилучший результат был получен на нейронной сети с конфигурацией №4. В ходе экспериментов было установлено, что нейронные сети с функцией активации relu показывают более точные результаты. Так же увеличить точность помогло увеличение нейронов на всех слоях. Однако, точность увеличилась не сильно, а время работы более чем в 3 раза.

# Литература

1. Образовательный курс «Методы глубокого обучения для решения задач компьютерного зрения», Кустикова В. Д.