

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное
образовательное учреждение высшего образования
Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского

Институт информационных технологий, математики и механики

Кафедра математического обеспечения и суперкомпьютерных технологий

Направление подготовки
02.04.02. Фундаментальная информатика и информационные технологии

Направленность образовательной программы
магистерская программа «Компьютерная графика и моделирование живых и
технических систем»

Отчёт
по методам глубокого обучения для решения задач компьютерного зрения

на тему:
«Начальная настройка весов полностью связанных нейронных сетей»

Квалификация (степень)
магистр

Форма обучения
очная

Выполнили: студенты группы 381706-3М
Храмов Илья Валерьевич

Подпись

Реунова Ольга Алексеевна

Подпись

Воеводин Андрей Михайлович

Подпись

Н. Новгород
2018 г.

Содержание

Постановка задачи.....	3
Тренировочные и тестовые наборы данных.....	4
Метрика качества решения.....	6
Разработанные программы/скрипты.....	7
Тестовые конфигурации сетей.....	8
Результаты.....	10
Литература.....	13

Постановка задачи

Цель

Цель настоящей работы состоит в том, чтобы использовать методы обучения без учителя для настройки начальных значений весов сетей, построенных при выполнении предшествующих практических работ.

Задачи

Выполнение лабораторной работы предполагает решение следующих задач:

1. Выбор архитектур нейронных сетей, построенных при выполнении предшествующих практических работ.
2. Выбор методов обучения без учителя для выполнения настройки начальных значений весов сетей.
3. Применение методов обучения без учителя к выбранному набору сетей.
4. Сбор результатов экспериментов.

Тренировочные и тестовые наборы данных

Задача — классификация комиксов. Данные получены из [5]. 86 классов сокращены до 14 классов с целью убрать классы, в которых выборка не репрезентативна. Размер изображений в каждом классе 288*432. Изображения 3 каналные.

№	Категории	Размер тренировочной выборки	Размер тестовой выборки
1	Aquaman v7	1088	282
2	Batgirl v4	1088	293
3	Batman v2	1509	384
4	Batwing	616	167
5	Batwoman	739	187
6	Catwoman v4	1047	256
7	Green Arrow	1040	256
8	Green Lantern	1301	345
9	Harley Quinn	671	165
10	Nightwing v3	558	147
11	Red Lanterns	767	197
12	Sinestro	619	143
13	Supergirl v6	701	200
14	Wonder Woman	1233	292
		12977	3314

Таблица 1. Размер выборки в каждом классе тренировочного и тестового множеств.

Изображения хранятся в формате JPEG.

На вход сети подаются бинарные файлы с расширениями .rec (изображения), .idx (индексы изображений) (Рис. 1).

Binary Record

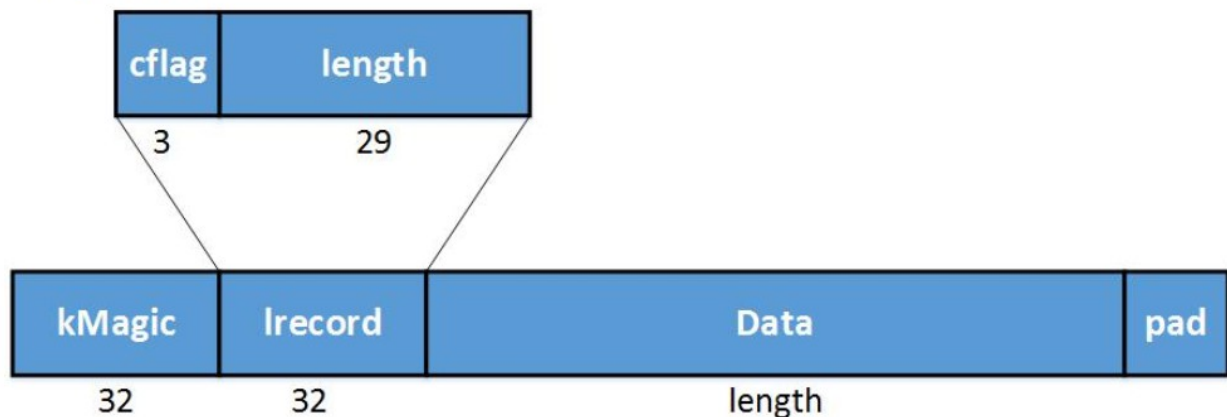


Рисунок 1. Формат хранения в MXNet [8].

- kMagic – начало записи;
- lrecord – длина (length) и продолжительность записи (cflag);
- Data – данные;
- pad – пространство для выравнивания до 4 байт.

Метрика качества решения

В качестве метрики для оценки качества решения задачи [5] выбрана “Точность” (“Accuracy”) и ”средняя абсолютная ошибка” (MAE). В терминологии MXNet, Accuracy — это отношение количества правильно предсказанных сэмплов к общему количеству сэмплов (Рис. 2).

$$\text{accuracy}(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} 1(\hat{y}_i == y_i) \qquad \text{MAE} = \frac{\sum_i^n |y_i - \hat{y}_i|}{n}$$

Рисунок 2. Определение “Accuracy” и MAE в MXNet [7].

Более подробную информацию о метриках можно найти в [7].

Разработанные программы/скрипты

- `reduce_dataset.py` - скрипт для вычленения из оригинального набора данных наиболее репрезентативных категорий.
- `prepare_dataset.py` - скрипт для подготовки данных под `mxnet`.
- `lab4.py` - скрипт для обучения автокодировщика.
- `resize_dataset.py` - скрипт для изменения размерности данных.
- `load_dataset.py` - функции для загрузки данных в скрипт обучения.
- `blocks.py` - основные конструкционные блоки сетей.
- `fit.py` - универсальный скрипт обучения, формирующий все необходимые отчёты.
- `parse_log.py` - скрипт, формирующий сводную таблицу из журнала обучения.

Тестовые конфигурации сетей

Последовательное обучение стека автокодировщиков представлено на рисунках 3-9. На рисунке 10 представлена полносвязная сеть прямого распространения с начальной настройкой весов.

1.

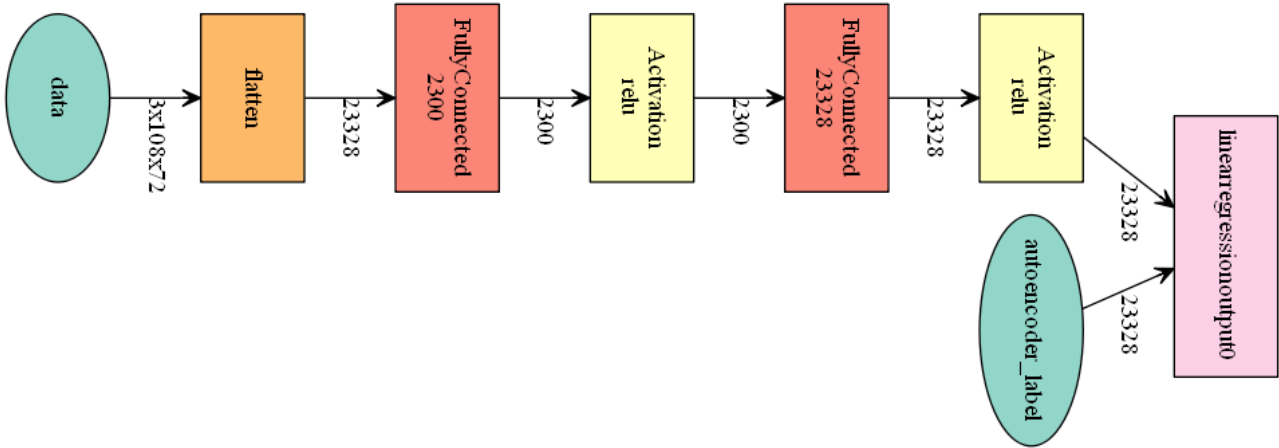


Рисунок 3. Обучение 1 автокодировщика (A1).

2.

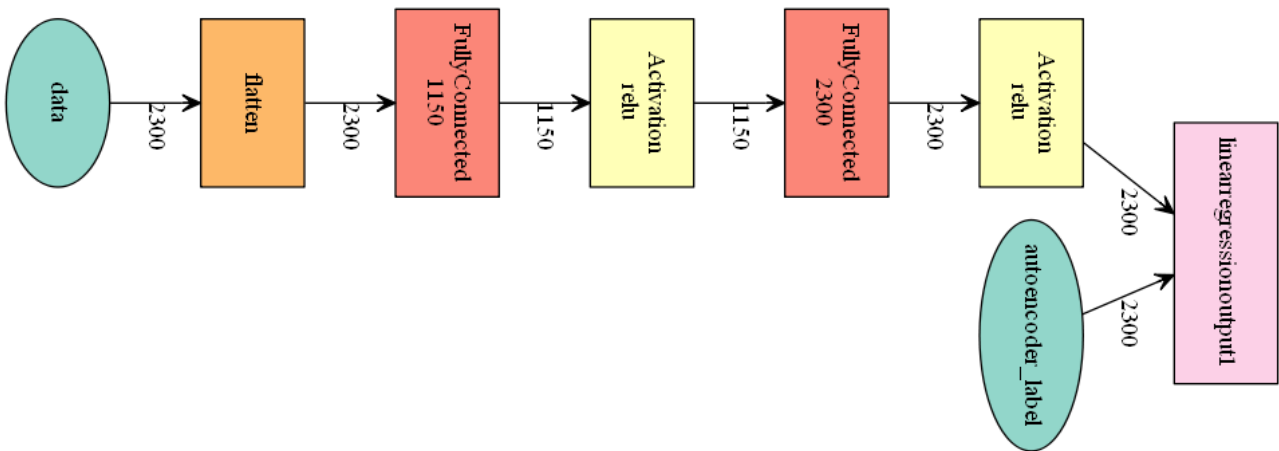


Рисунок 4. Обучение 2 автокодировщика (A2).

3.

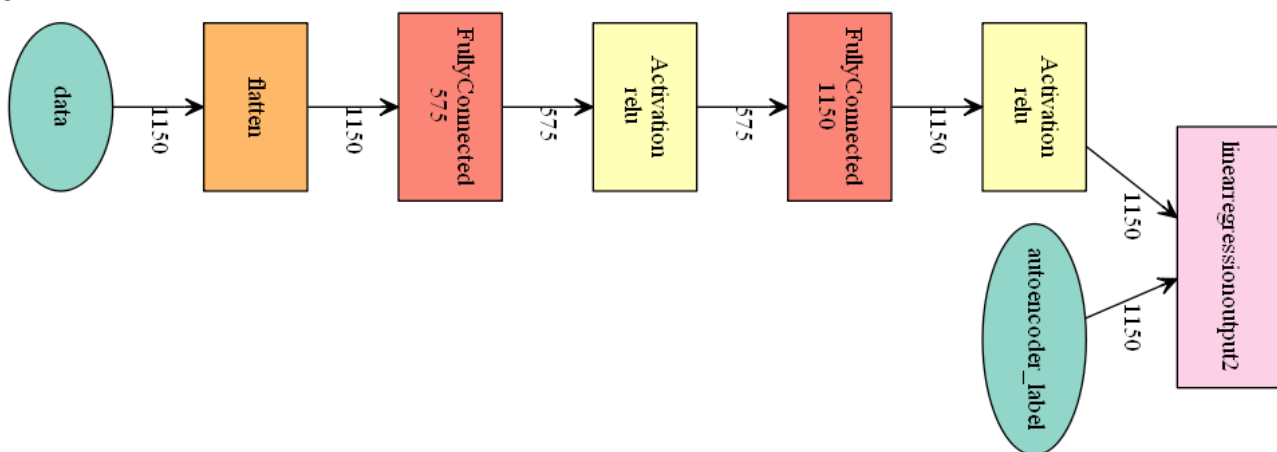


Рисунок 5. Обучение 3 автокодировщика (A3).

4.

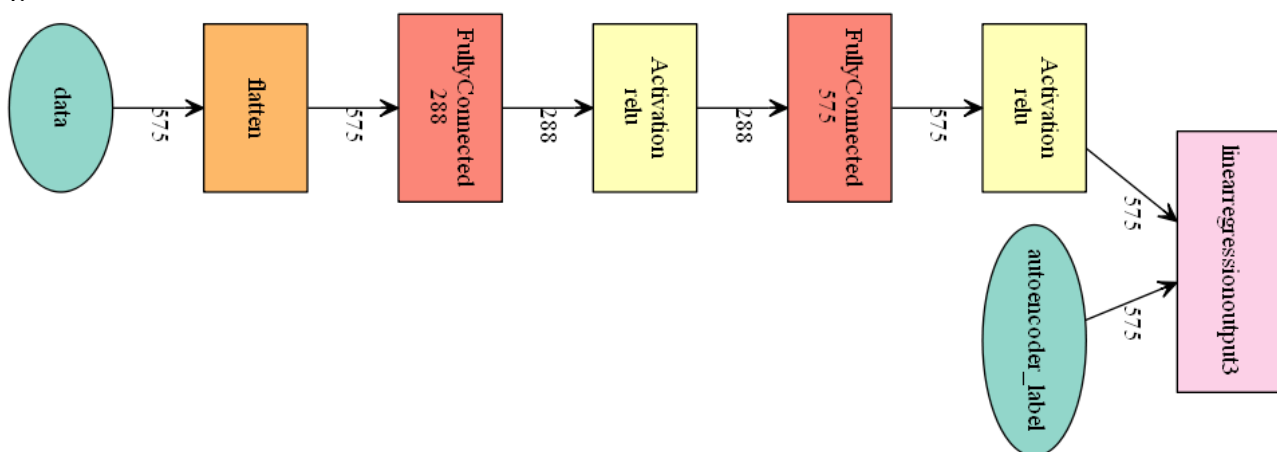


Рисунок 6. Обучение 4 автокодировщика (A4).

5.

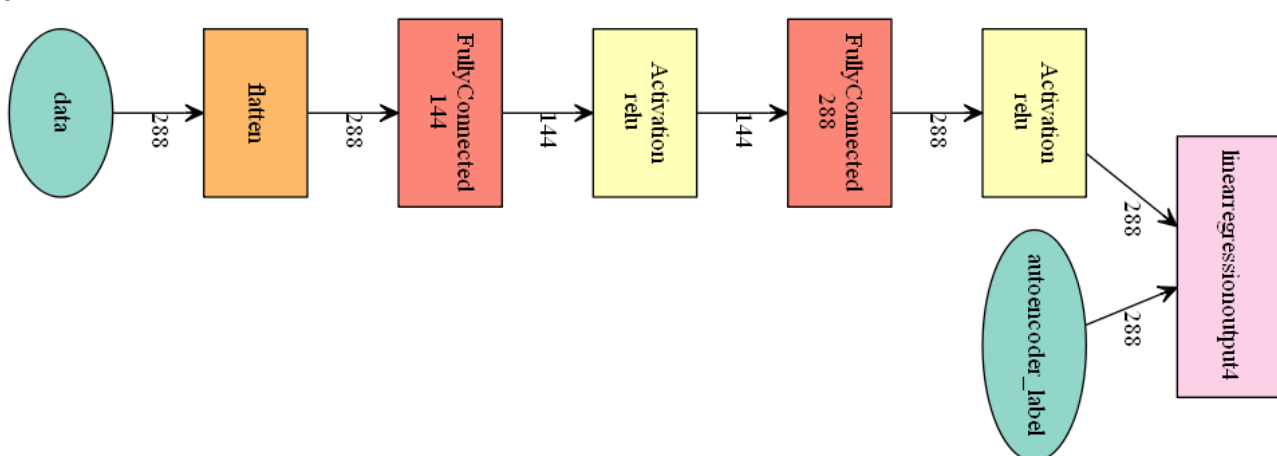


Рисунок 7. Обучение 5 автокодировщика (A5).

6.

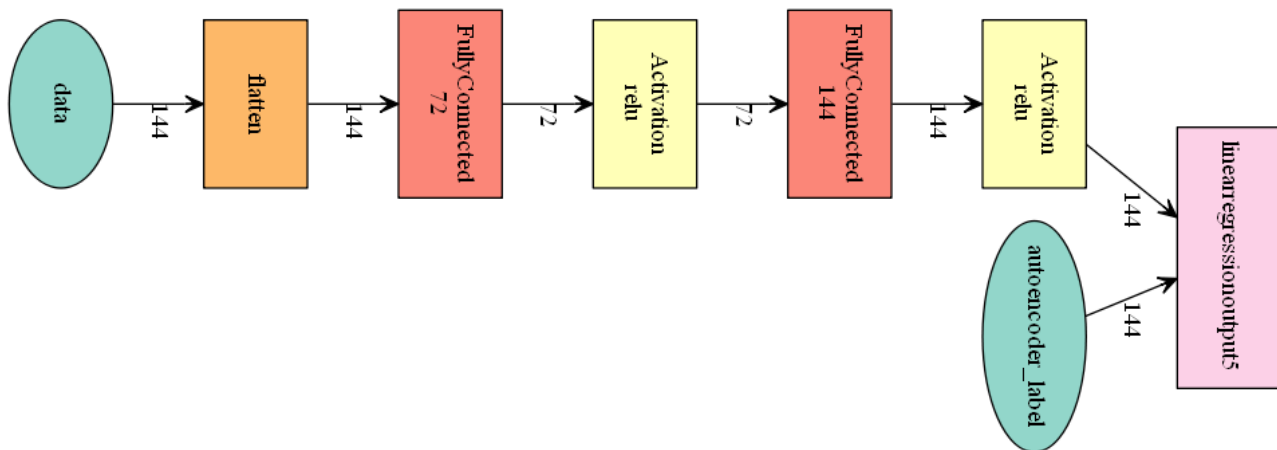


Рисунок 8. Обучение 6 автокодировщика (A6).

7.

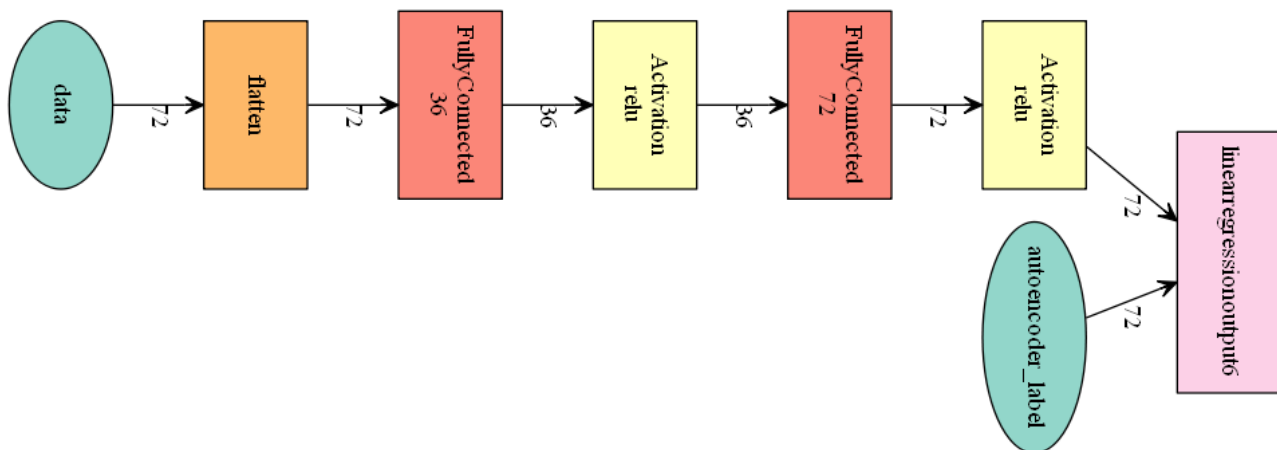


Рисунок 9. Обучение 7 автокодировщика (A7).

8.

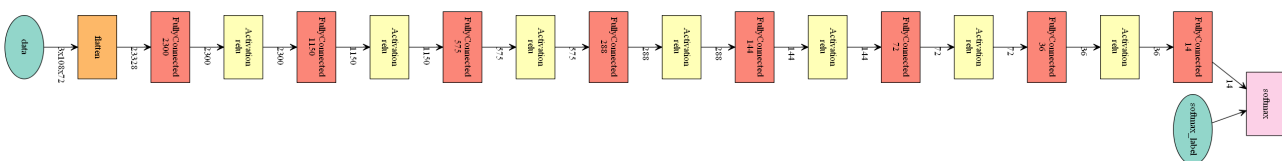


Рисунок 10. Сеть1 с начальной настройкой весов.

Результаты

В таблице 2 приведены конфигурация системы и программное обеспечение, с помощью которых проводилось обучение и тестирование построенных моделей.

Параметры	Версия
Операционная система	Windows10
GPU	NVIDIA GeForce GTX 1080 (частота процессора — OC Mode – GPU Boost Clock : 1835 MHz , GPU Base Clock : 1695 MHz Gaming Mode (Default) - GPU Boost Clock : 1809 MHz , GPU Base Clock : 1670 MHz; шина передачи данных – PCI Express 3.0; видеопамять - GDDR5X 8GB; количество ядер — 2560; частота памяти - 10010 MHz; интерфейс памяти — 256-bit;)
CUDA	9.2
Python	3.7.1
MXNet	1.3.0

Таблица 2. Конфигурация системы.

В таблице 3 приведены параметры обучения.

Наименование сети	Оптимизатор	Скорость обучения	Количество эпох	Размер batch
Автокодировщик1 (A1)	SGD	0.0001	15	10
Автокодировщик2 (A2)	SGD	0.0001	15	10
Автокодировщик3 (A3)	SGD	0.0001	15	10
Автокодировщик4 (A4)	SGD	0.0001	15	10
Автокодировщик5 (A5)	SGD	0.0001	15	10
Автокодировщик6 (A6)	SGD	0.0001	15	10
Автокодировщик7 (A7)	SGD	0.0001	15	10
Сеть1	SGD	0.0001	60	10

Таблица 3. Параметры обучения.

	Наименование сети
	Сеть1
Среднее время обучения за одну эпоху, с	10,46
Качество решения на тренировочном наборе (Assurasy), %	11,13
Качество решения на тестовом наборе (Assurasy), %	11,54
Номер эпохи с достигнутым максимальным качеством решения на тренировочном наборе	24
Максимальное качество решения на тренировочном наборе (Assurasy), %	12,47
Номер эпохи с достигнутым максимальным качеством решения на тестовом наборе	43
Максимальное качество решения на тестовом наборе (Assurasy), %	11,63

Таблица 4. Результаты экспериментов. Конфигурация сетей приведена в “Тестовые конфигурации сетей”.

Наименование автокодировщика	Среднее время обучения за одну эпоху, с	Train-mae	Номер эпохи с достигнутым минимальным MAE	Минимальное MAE
A1	18,11	19876.825729	5	19202.914073
A2	3,89	0	1	0
A3	3,72	0	1	0
A4	3,4	0	1	0
A5	3,73	0	1	0
A6	3,76	0	3	0
A7	3,81	0	1	0

Таблица 5. Результаты экспериментов. Конфигурация автокодировщиков приведена в “Тестовые конфигурации сетей”.

Анализ результатов

Из представленных результатов можно сделать выводы:

1. Начальная настройка весов не помогла улучшить точность на тренировочной выборке (было 21,69).
2. Выборка — нерепрезентативная. Малое количество изображений на один класс, как тренировочной, так и тестовой выборки. Некоторые изображения в тестовой и тренировочной выборках слишком отличаются (например, тренировочная выборка — нарисован Бэтман на черном фоне, тестовая выборка — нарисован Бэтман на светлом

фоне). Герои одних комиксов могут встречаться в других комиксах (например, Harley Quinn в Batman и наоборот).

3. Плохой выбор скорости обучения. Если данный параметр слишком большой, то точность будет приблизительно 11% (измерялось для разных конфигураций сетей), то есть мы будем бесконечно «прыгать» через точку минимума. Если мы берем параметр слишком маленьким, то мы «застреваем» в локальном минимуме.
4. Начальная обработка данных. Для конкретного размера изображения необходимо настраивать количество нейронов так, чтобы избежать «узкого горлышка».
5. Надо понимать, что для таких задач лучше использовать сверточные или другие нейронные сети.

Литература

1. MNIST dataset [<http://yann.lecun.com/exdb/mnist>].
2. OpenCV [<http://opencv.org>].
3. Материалы Летней межвузовской школы 2016 [<https://github.com/itseez-academy/itseez-ss-2016-theory>], [<https://github.com/itseez-academy/itseez-ss-2016-practice>].
4. Лекции по глубокому обучению: <https://sites.google.com/site/kustikovavalentina/studentam/kurs-glubokoe-obucenie>, 2018.
5. Исходные данные — <https://www.kaggle.com/cenkbircanoglu/comic-books-classification>: kaggle datasets download -d cenkbircanoglu/comic-books-classification.
6. Документация MXNet — <http://mxnet.incubator.apache.org/test/tutorials/>.
7. Метрики в MXNet – <https://mxnet.incubator.apache.org/api/python/metric/metric.html>.
8. Формат хранения данных в MXNet: https://mxnet.incubator.apache.org/architecture/note_data_loading.html.
9. Репозиторий исходных кодов: <https://github.com/okondratieva/DeepLearning>.