Университет ИТМО

Факультет программной инженерии и компьютерной техники

Системы искусственного интеллекта

Лабораторная работа №4

### Вариант 3

Выполнила: Голованова Д.В.

Группа P33671

Проверил: Александр Владимирович Кугаевских

Санкт-Петербург,

2022 г.

# Задание

|  |  |  |  |
| --- | --- | --- | --- |
| Var | Part1 func | Part2 data | Hyperparameters |
| 1 | Absolute(Sin(x)) X: 6,3..6.3 Y: 0..1.2 | CIFAR10 | Layers count, neurons count per layer |
| 2 | Cos(x) X: -9..9 Y: -1..1 | CIFAR100 | Learn rate, regularization L1 |
| 3 | Absolute(Sin(x)) X: 6,3..6.3 Y: 0..1.2 | Handwritten digits | Regularization L2, output layer activation type |
| 4 | Cos(x) X: -9..9 Y: -1..1 | Fashion articles | Layer activation type, loss function type |

# Выполнение

Оптимизация гиперпараметров — задача машинного обучение по выбору набора оптимальных параметров для обучающего алгоритма. Одни и те же виды моделей машинного обучения могут требовать различные предположения, веса или скорости обучения для различных видов данных. Эти параметры называются гиперпараметрами и их следует настраивать так, чтобы модель могла оптимально решить задачу обучения.

## Часть 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| L2 | softmax | relu | tanh | | sigimod | linear |
| -0.1000 | 0.790 | 0.209 | 0.790 | | 0.819 | 0.209 |
| 0.0000 | 0.790 | 0.875 | 0.884 | | 0.790 | 0.865 |
| 0.0001 | 0.790 | 0.834 | 0.209 | | 0.875 | 0.834 |
| 0.0005 | 0.790 | 0.209 | 0.845 | | 0.824 | 0.209 |
| 0.0010 | 0.790 | 0.790 | 0.209 | | 0.790 | 0.209 |
| 0.0050 | 0.790 | 0.790 | 0.860 | | 0.790 | 0.814 |
| 0.0100 | 0.790 | 0.209 | 0.860 | | 0.839 | 0.790 |
| 0.0500 | 0.790 | 0.790 | 0.824 | | 0.790 | 0.800 |
| 0.1000 | 0.790 | 0.7900 | 0.790 | | 0.790 | 0.209 |
|  | | | | | | |
| Лучшие найденные параметры:   * Функция активации — tanh * L2-регуляризация — 0.0000 * Полученная точность — 88.4% | | | |  | | |
|  | | | | | | |

## Часть 2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| L2 | softmax | relu | tanh | sigimod | linear |
| -0.1000 | 0.113 | 0.098 | 0.098 | 0.113 | 0.101 |
| 0.0000 | 0.113 | 0.098 | 0.098 | 0.113 | 0.098 |
| 0.0001 | 0.113 | 0.098 | 0.098 | 0.113 | 0.098 |
| 0.0005 | 0.113 | 0.098 | 0.098 | 0.113 | 0.098 |
| 0.0010 | 0.113 | 0.098 | 0.098 | 0.113 | 0.098 |
| 0.0050 | 0.113 | 0.098 | 0.098 | 0.276 | 0.089 |
| 0.0100 | 0.113 | 0.098 | 0.098 | 0.298 | 0.098 |
| 0.0500 | 0.113 | 0.185 | 0.098 | 0.802 | 0.098 |
| 0.1000 | 0.113 | 0.098 | 0.098 | 0.251 | 0.098 |

Лучшие найденные параметры:

* Функция активации — sigmoid
* L2-регуляризация — 0.05
* Полученная точность — 80.2%

При обучении 4 слоями по 20 нейронов в течение 100 эпох точность достигает 93.8% (0.9381999969482422)

# Вывод

В процессе выполнения лабораторной, я немного «поигралась» с настоящей нейросетью, с целью оптимизировать гиперпараметры для улучшения результатов работы сети на наборе данных из библиотеки mnist. Эта лабораторная является интересной точкой старта в data science, а также увидела, как влияют некоторые параметры на результаты (к примеру, LR не может быть отрицательным). Для выполнения использовала бесплатный сервис Google Colab.