**Тема №1: Реализация многослойного перцептрона.**

**Цель:** реализовать многослойный перцептрон с различными функциями активации

**Особенности реализации:** torch.nn

**Эксперимент:** реализовала многослойный перцептрон со следующими функциями активации:

* Сигмоида
* Гиперболический тангенс
* ReLU

Количество слоев варьируется от 2 до 4, количество нейронов – от 1 до 5

Код реализации: <https://colab.research.google.com/drive/1g17VE_i9zSIcxJkSNaHlLHriyRlxJHJP?usp=sharing>

**Тема №2: Размер выборки, необходимой для получения классификатора с необходимой точностью. Кросс-валидация.**

**Цель:** определить размер выборки, необходимый для получения классификатора с заданной точностью и провести кросс-валидацию

**Особенности реализации:** torch.nn

**Эксперимент:**

1. Сначала я оценила количество элементов в обучающей выборке, достаточное для получения классификатора, позволяющего прогнозировать значения y с точностью не ниже 90%. Я зафиксировала количество эпох и learning rate, и так же рассмотрела случаи различных пар значений количества слоев и нейронов для всех трех функций активации
2. Далее я провела кросс-валидацию для того, чтобы определить наилучшие значения функции активации и количества слоев

**Результаты:**

**Линейная выборка**

Epochs: 100

Learning rate: 0.01

Размер обучающей выборки: 500

Размер тестовой выборки: 200

* сигмоида

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 | 1.0 | 1.0 | 1.0 | 1.0 |
| 3 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| 4 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

* тангенс

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| 3 | 1.0 | 1.0 |  |  |  |
| 4 |  |  |  |  | 1.0 |

* ReLU

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 | 1.0 |  |  | 1.0 |
| 3 | 0.5 | 0.5 |  |  |  |
| 4 | 0.5 |  |  |  | 1.0 |

**Circles**

Epochs: 100

Learning rate: 0.01

Размер обучающей выборки: 500

Размер тестовой выборки: 200

* сигмоида

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| 3 | 0.5 | 0.5 |  |  |  |
| 4 | 0.5 |  |  | 0.6 | 0.5 |

* тангенс

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.649 | 0.819 | 1.0 | 1.0 | 1.0 |
| 3 | 0.5 | 0.8 | 1.0 | 1.0 | 1.0 |
| 4 | 0.5 | 0.65 | 0.5 | 0.8248 | 0.639 |

* ReLU

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 | 0.5 | 1.0 |  |  |
| 3 | 0.5 | 0.5 | 0.66 | 0.995 |  |
| 4 | 0.5 | 0.5 |  |  | 1.0 |

**Spiral**

Epochs: 100

Learning rate: 0.01

Размер обучающей выборки: 500

Размер тестовой выборки: 200

* сигмоида

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  | 0.5 |

* тангенс

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.54 | 0.55 |  |  | 0.575 |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  | 0.56 |

Лучший результат с тангенсом для спирали, которого я смогла добиться – 0.8, параметры этого эксперимента:

размер обучающей выборки – 70, тестовой – 10, количество эпох – 500, 4 слоя, 5 нейронов в каждом

* ReLU

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 |  |  |  |  |
| 3 |  |  |  |  | 0.5 |
| 4 |  |  | 0.52 | 0.52 | 0.51 |

**XOR**

Epochs: 100

Learning rate: 0.01

Размер обучающей выборки: 500

Размер тестовой выборки: 200

* сигмоида

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 | 0.5 |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  | 0.5 |

* тангенс

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 | 0.75 | 0.995 |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  | 0.995 |

* ReLU

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Слои/нейроны | 1 | 2 | 3 | 4 | 5 |
| 2 | 0.5 |  |  |  |  |
| 3 |  |  |  |  |  |
| 4 |  |  |  |  | 0.995 |

**Кросс-валидация**

Размер общей выборки: 700

Количество подмножеств, на которые делится выборка: 4

Количество эпох: 100

Learning rate: 0.01

Количество нейронов в каждом эксперименте: 5

**Линейная выборка**

* Сигмоида

2 слоя, среднее значение: 0.0028

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя, среднее значение: 0.032

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* Тангенс

2 слоя, среднее значение: 0.00024

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя, среднее значение: 0.00023

Изображение выглядит как текст, Шрифт, снимок экрана, черно-белый

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя, среднее значение: 0.00021

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* ReLU

2 слоя, среднее значение: 0.00032

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя, среднее значение: 0.000167

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя, среднее значение: 0.0027

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

**Circles**

* Сигмоида

2 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* Тангенс

2 слоя, среднее значение: 0.0121

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя: среднее значение: 0.00115

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя, среднее значение: 0.0038

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* ReLU

2 слоя, среднее значение: 0.0123

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

**Spiral**

* Сигмоида

2 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* Тангенс

2 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* ReLU

2 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

**XOR**

* Сигмоида

2 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* Тангенс

2 слоя, среднее значение: 0.596

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

3 слоя, среднее значение: 0.0319

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя, среднее значение: 0.046

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

* ReLU

2 слоя, среднее значение: 0.1835

**Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.**

3 слоя, среднее значение: 0.04

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

4 слоя, среднее значение: 0.079

Изображение выглядит как текст, Шрифт, снимок экрана

Контент, сгенерированный ИИ, может содержать ошибки.

Вывод:

1. Многослойный перцептрон намного лучше справляется с любым типом выборок, чем однослойный перцептрон. С обучающей выборкой размера 500 и тестовой размера 200 удалось достичь заданной точности практически для всех случаев, кроме выборки «спиралью». Так же с функцией активацией сигмоиды было сложнее достичь точности, так как с ней обучение происходит намного медленнее, чем с тангенсом или ReLU

Для сигмоиды мне не удалось подобрать параметры (за исключением случая линейной выборки), при которых точность была бы не меньше 90%

1. В результате проведения кросс-валидации я подобрала наилучшие значения параметров функции активации и количества слоев для выборок размера 700 (количество нейронов - 5)

Линейная выборка: ReLU 3 слоя

Circles: тангенс 3 слоя

XOR: тангенс 3 слоя

Что касается спирали, сложно сказать, какая из функций активации подходит лучше, так как значение ошибки велико во всех случаях

Так же можно отметить, что наблюдается та же тенденция: перцептрон с сигмоидой обучается хуже, поэтому она не особо эффективна