Университет ИТМО
Факультет программной инженерии и компьютерной техники

Курс «Искусственный интеллект»

Лабораторная работа № 4

Вариант: 1

Работу выполнил

Хуан Сыюань
P33101

Преподаватель
Игорь Александрович
Бессмертный

Санкт-Петербург
2021

Оглавление

Цель лабораторной работы

Описание предметной области

Study of the hyperparameters of the neural network, understanding impact of different hyperparameters on accuracy.

Задание

Part 1 represent recognition of basic math functions with illustrations of neural net vision of the functions for training.

Part 2 represent recognition of simple images for making following work.

1) By changing these hyperparameters try to reach max accuracy value(at least 0.95) for Part2 model with fixed epoch count 20

2) Change 1st hyperparameter’s value from min to max with minimal step depends on your variant

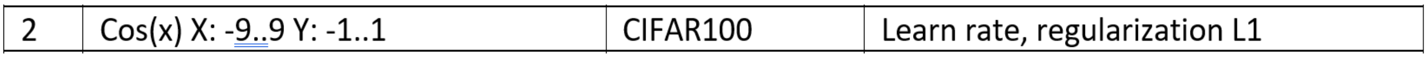
3) Show impact on result using graphs

4) Describe impact of each hyperparameter on accuracy.

5) Set hyperparameter value back to one which produced max accuracy

6) Repeat 2-5 steps for second hyperparameter

Вариант



2	Cos(x) X: -9..9 Y: -1..1	CIFAR100	Learn rate, regularization L1

Выполнение лабораторной работы

Описание параметров

Learn rate (Коэффициент скорости обучение)

This parameter is the amount of weight correction at each iteration.
It is in the range from 0 to 1 (0 is not included, since this value will not correct the balance at all).

The larger the correction step, the faster the algorithm works.
On the other hand, a large step increases the training error, since the accuracy of the model tuning decreases to a minimum of the error function.

Experimental conditions:
We will gradually increase the learning rate factor by running the model at each of the values.
In this case, the speed of the algorithm will not be taken into account.

Based on these conditions, we can conclude that the training accuracy should decrease all the time.

Regularization L1 (Регуляризация L1)

To understand what regularization is for, it is necessary to introduce two concepts:
Overfitting is a negative phenomenon that occurs when a learning algorithm generates predictions that match too closely or exactly a specific set of data and therefore are not suitable for applying the algorithm to additional data or future observations.
Underfitting is a negative phenomenon in which the learning algorithm does not provide a sufficiently small value of the average error on the training sample. Underlearning occurs when insufficiently complex models are used. Therefore, in practical applications, we need to adjust the parameters to avoid these two situations.
When choosing a complex model with a lack of data, it is possible to obtain a final model that describes the training set well, but does not generalize to the test set.
One of the ways to deal with the negative effect of overfitting the data is to use regularization, that is, to add some penalty for large values of the coefficients of the linear model. This prevents too "sharp" bends and prevents overfitting.

Overfitting in most cases manifests itself in the fact that the resulting models have too large parameter values. Accordingly, it is necessary to add a penalty for this to the objective function.

Regularization L1 (Lasso regression) adds the absolute value of the coefficient value as a penalty to the loss function.

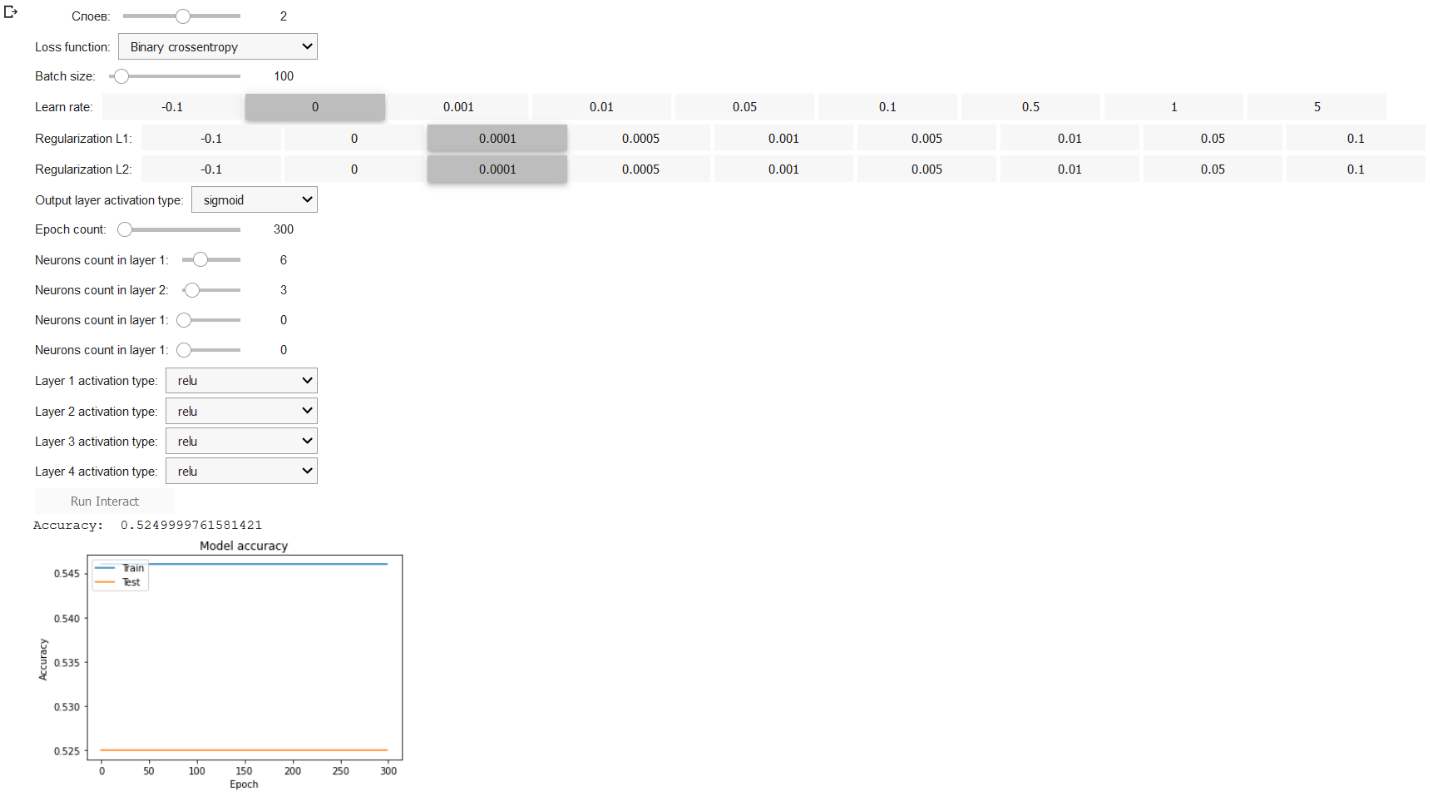
$$\sum_{i=1}^{n}(Y_i - \sum_{j=1}^{p} X_{ij}\beta_j)^2 + \lambda\sum_{j=1}^{p}|\beta_j|$$

Again, if lambda is zero then we will get back OLS whereas very large value will make coefficients zero hence it will under-fit.
If the lambda is zero, then we will get back the least squares method, while a large value will make the coefficients equal to zero, which will result in “under-fit”.
The use of the L1 regularization is that it brings the coefficients of the least important features closer to zero. This results in a selection of features, which works well when there are a lot of them.

Часть 1

Parameter change “Learn rate”

Learn rate: 0



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | 0.1 | 0.5 | 1 | 5

Regularization L1: -0.1 | 0 | 0.0001 | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | 0.0001 | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

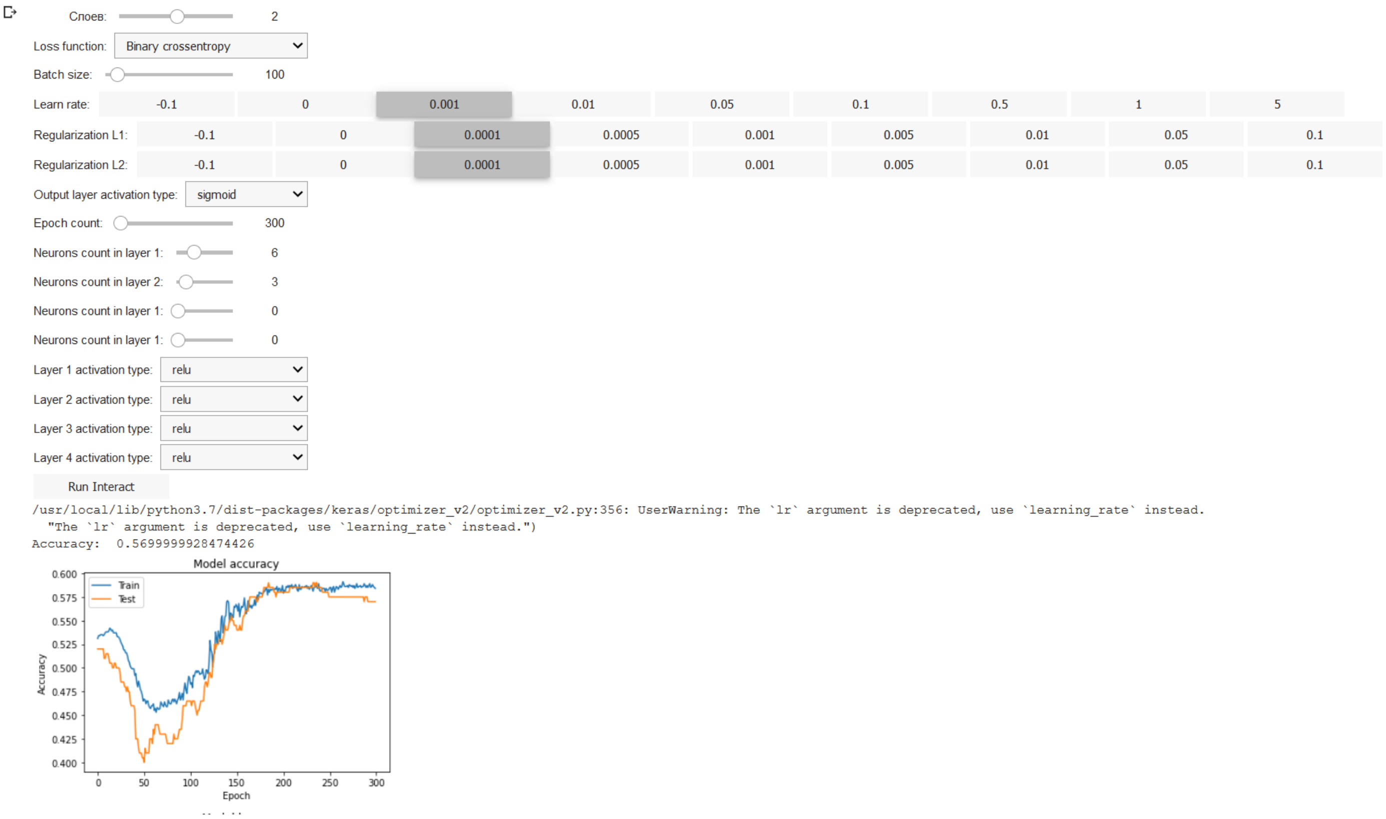
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.5249999761581421
```

Learn rate: 0.001



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | [✓ 0.001] | 0.01 | 0.05 | 0.1 | 0.5 | 1 | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

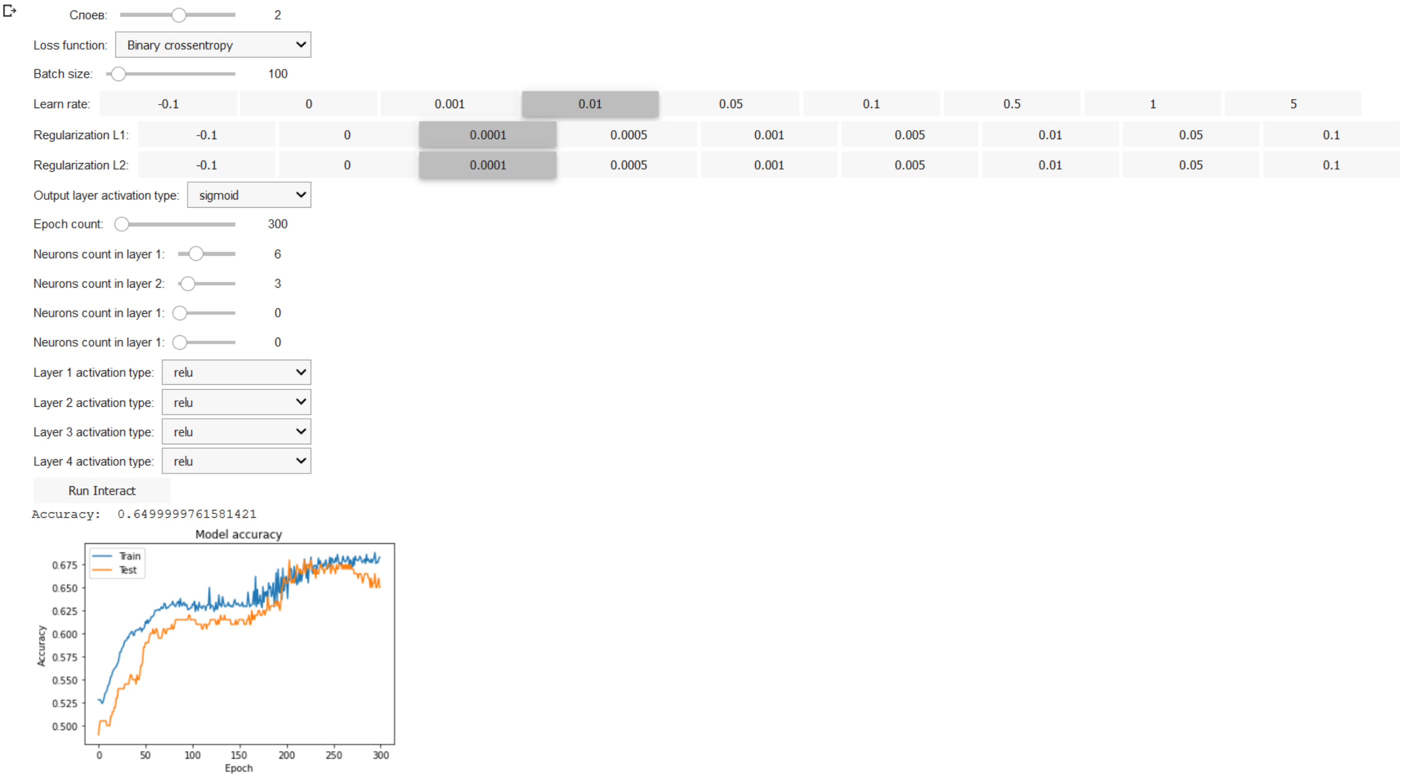
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
/usr/local/lib/python3.7/dist-packages/keras/optimizer_v2/optimizer_v2.py:356: UserWarning: The `lr` argument is deprecated, use `learning_rate` instead.
  "The `lr` argument is deprecated, use `learning_rate` instead.")
Accuracy:  0.5699999928474426
```

Learn rate: 0.01



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | [✓ 0.01] | 0.05 | 0.1 | 0.5 | 1 | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

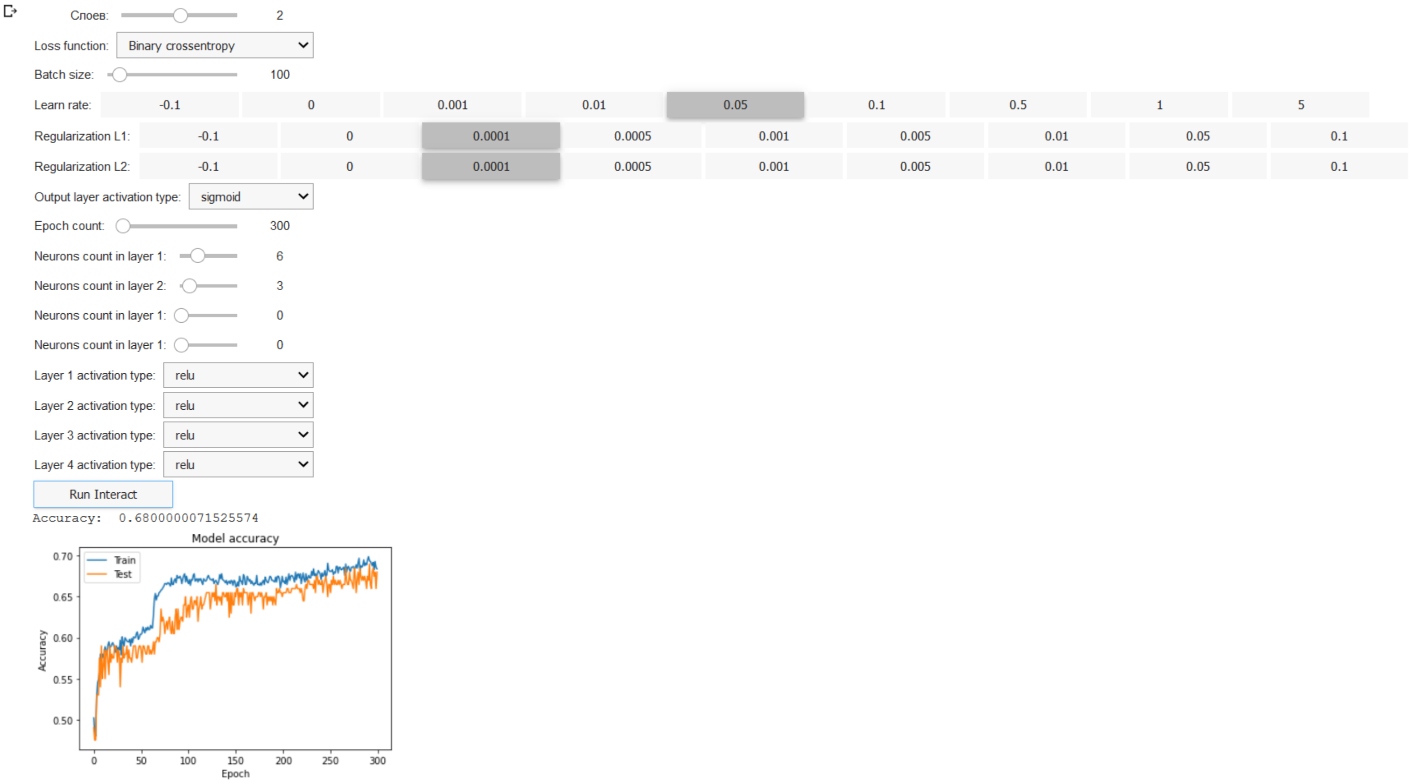
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.6499999761581421
```

Learn rate: 0.05



Cnoeв: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate:	-0.1	0	0.001	0.01	[✓ 0.05]	0.1	0.5	1	5
Regularization L1:	-0.1	0	[✓ 0.0001]	0.0005	0.001	0.005	0.01	0.05	0.1
Regularization L2:	-0.1	0	[✓ 0.0001]	0.0005	0.001	0.005	0.01	0.05	0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

Layer 2 activation type: relu

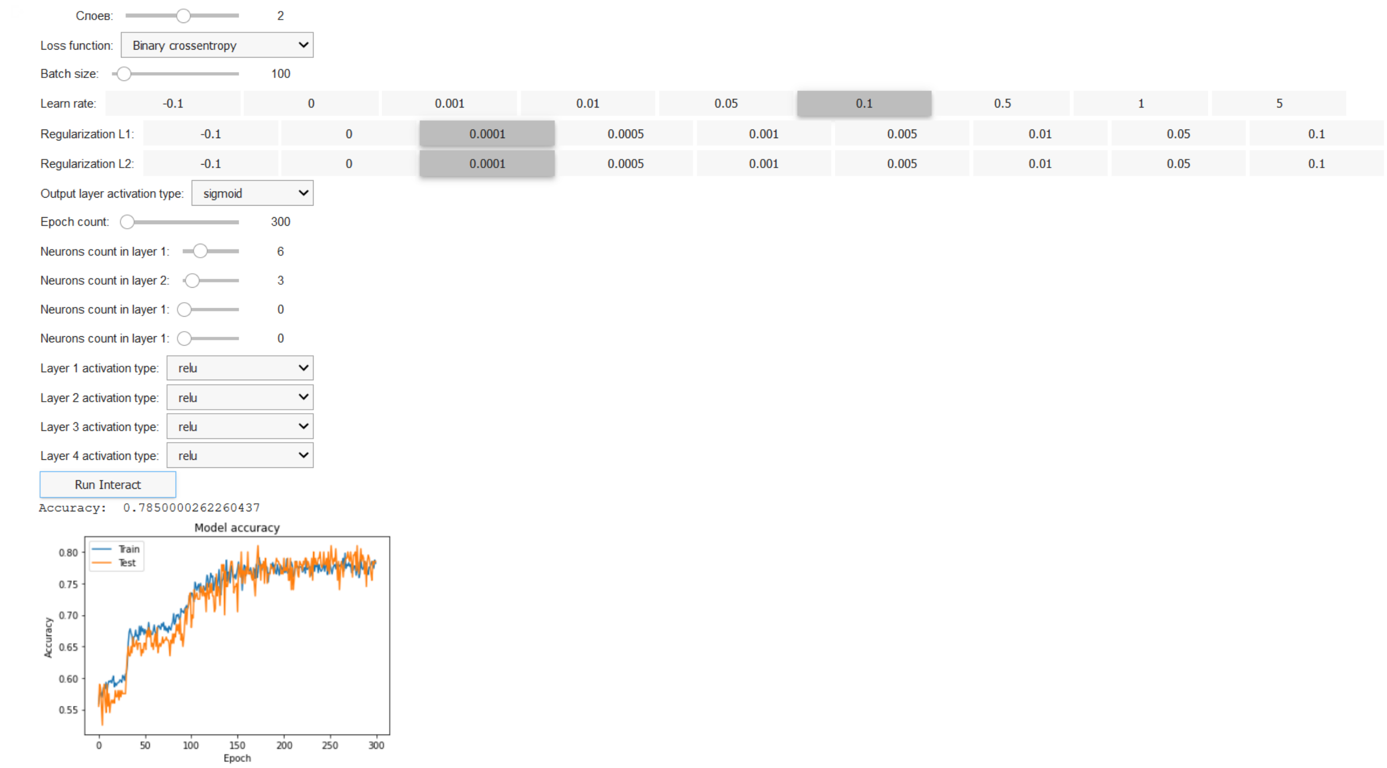
Layer 3 activation type: relu

Layer 4 activation type: relu

Run Interact

```
Accuracy:  0.6800000071525574
```

Learn rate: 0.1



Cnoeв: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate:	-0.1	0	0.001	0.01	0.05	[✓ 0.1]	0.5	1	5
Regularization L1:	-0.1	0	[✓ 0.0001]	0.0005	0.001	0.005	0.01	0.05	0.1
Regularization L2:	-0.1	0	[✓ 0.0001]	0.0005	0.001	0.005	0.01	0.05	0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

Layer 2 activation type: relu

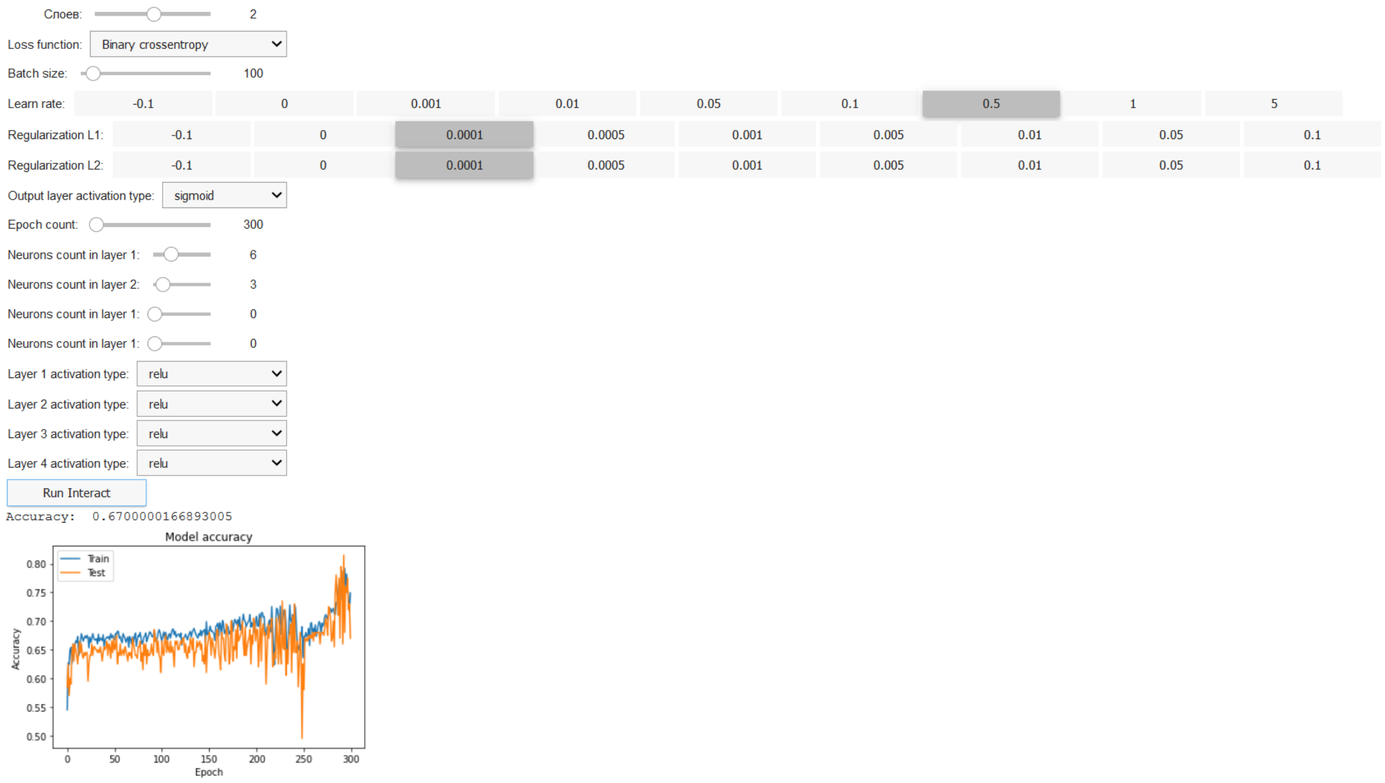
Layer 3 activation type: relu

Layer 4 activation type: relu

Run Interact

```
Accuracy:  0.7850000262260437
```

Learn rate: 0.5



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | 0.1 | [✓ 0.5] | 1 | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

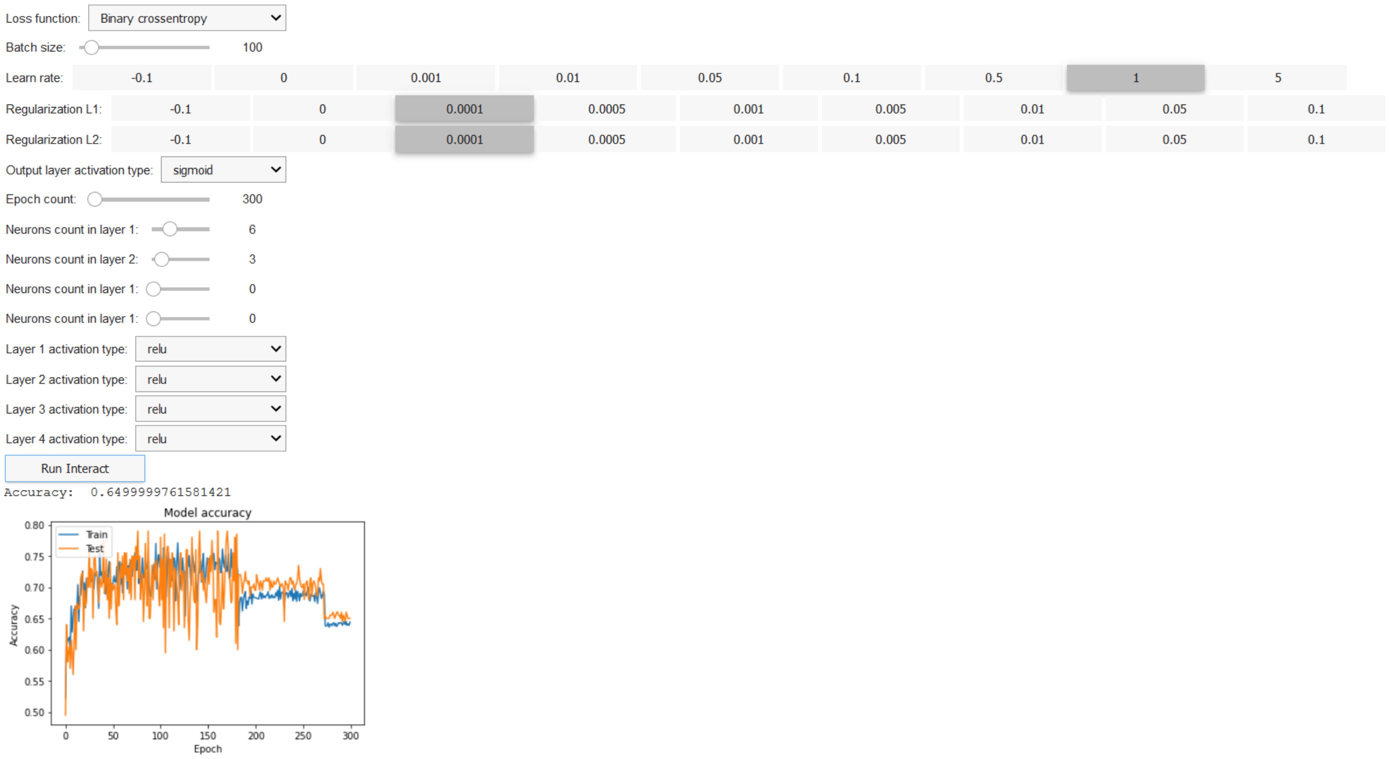
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.6700000166893005
```

Learn rate: 1



Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | 0.1 | 0.5 | [✓ 1] | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

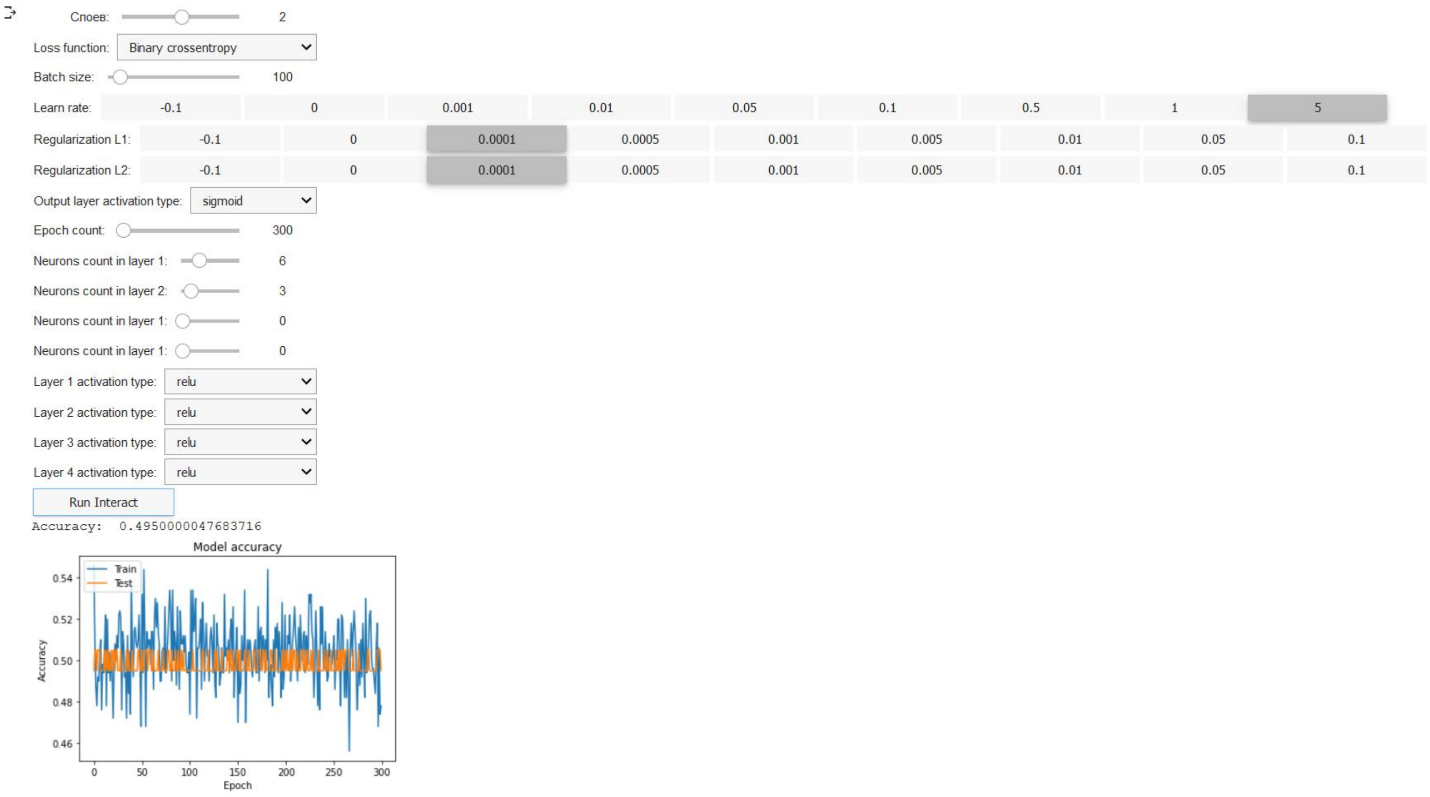
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.6499999761581421
```

Learn rate: 5



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | 0.1 | 0.5 | 1 | [✓ 5]

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

Layer 2 activation type: relu

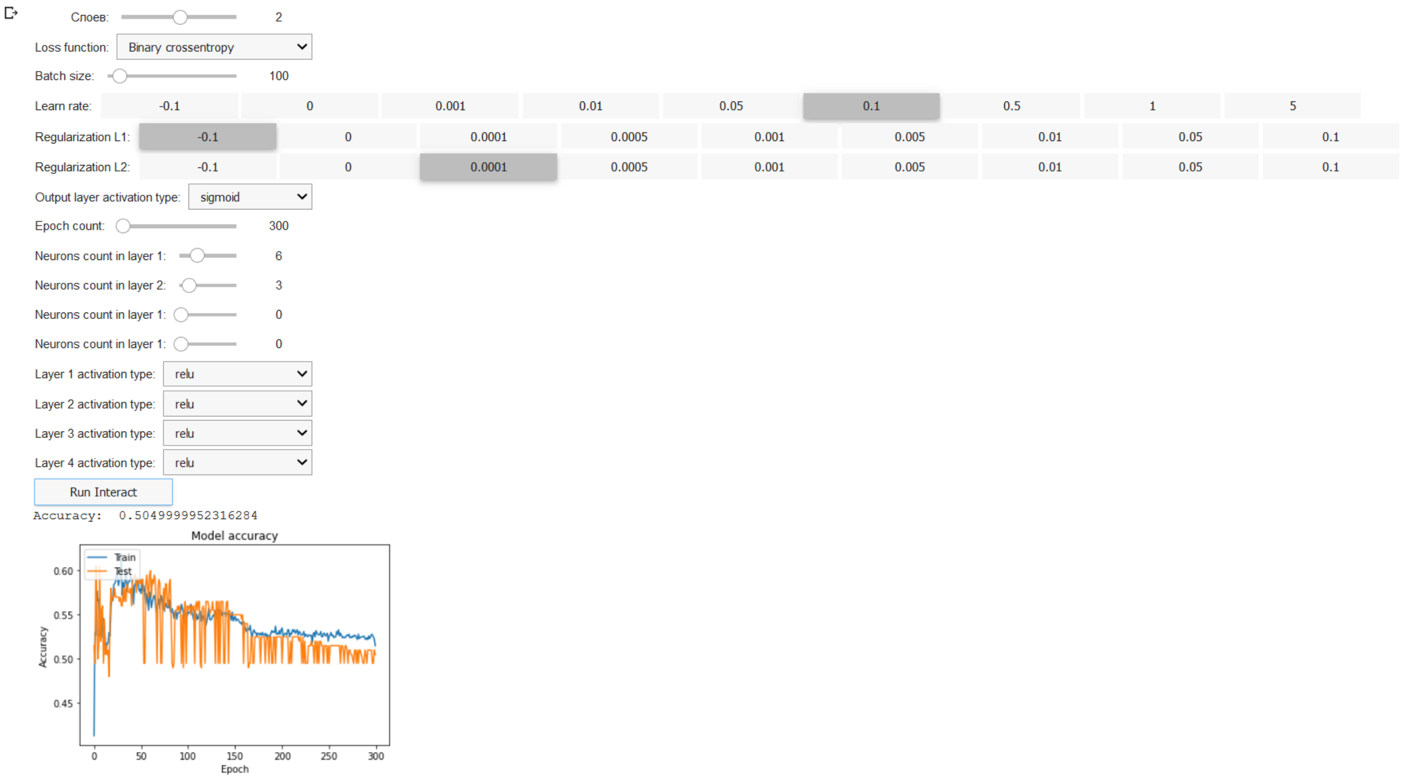
Layer 3 activation type: relu

Layer 4 activation type: relu

Run Interact

```
Accuracy:  0.4950000047683716
```

Parameter change “Regularization L1”



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | [✓ 0.1] | 0.5 | 1 | 5

Regularization L1: [✓ -0.1] | 0 | 0.0001 | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

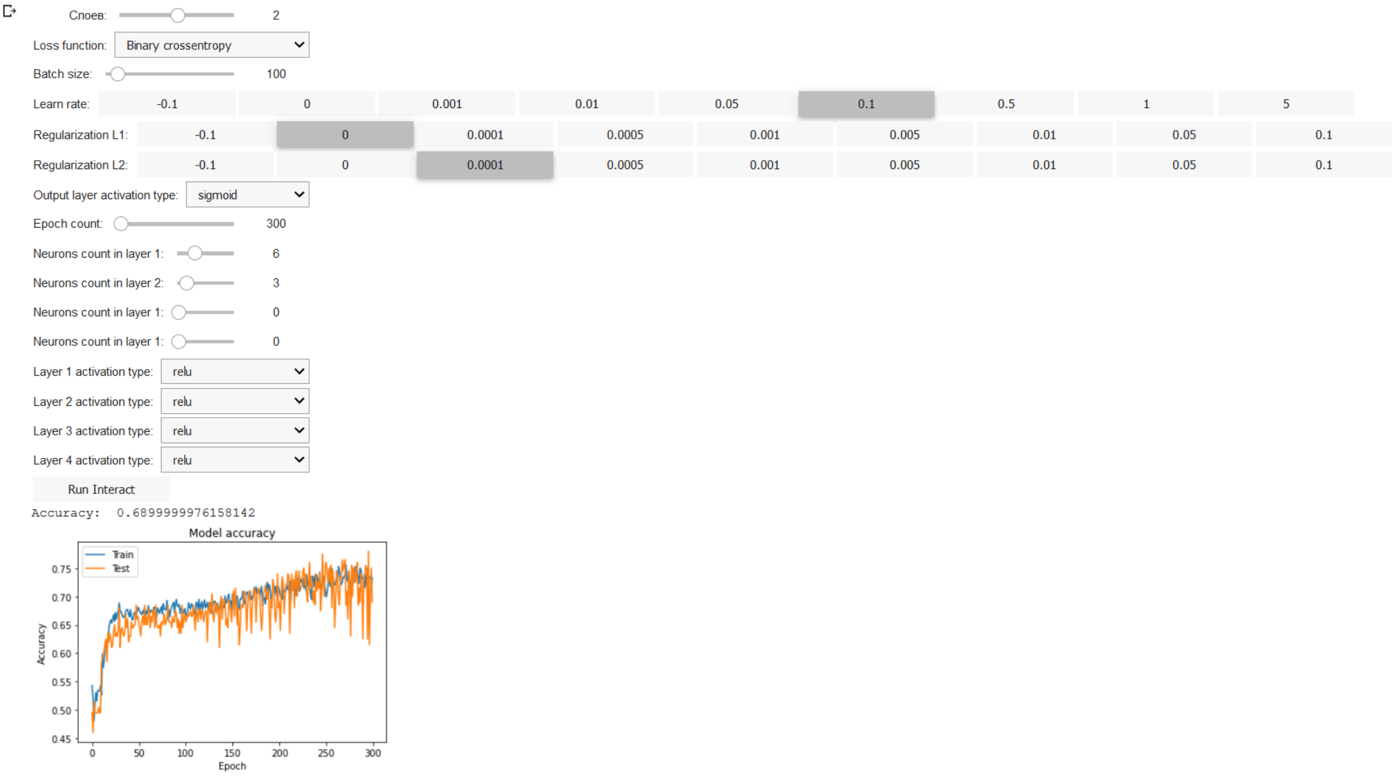
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

Run Interact

```
Accuracy:  0.5049999952316284
```



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | [✓ 0.1] | 0.5 | 1 | 5

Regularization L1: -0.1 | [✓ 0] | 0.0001 | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

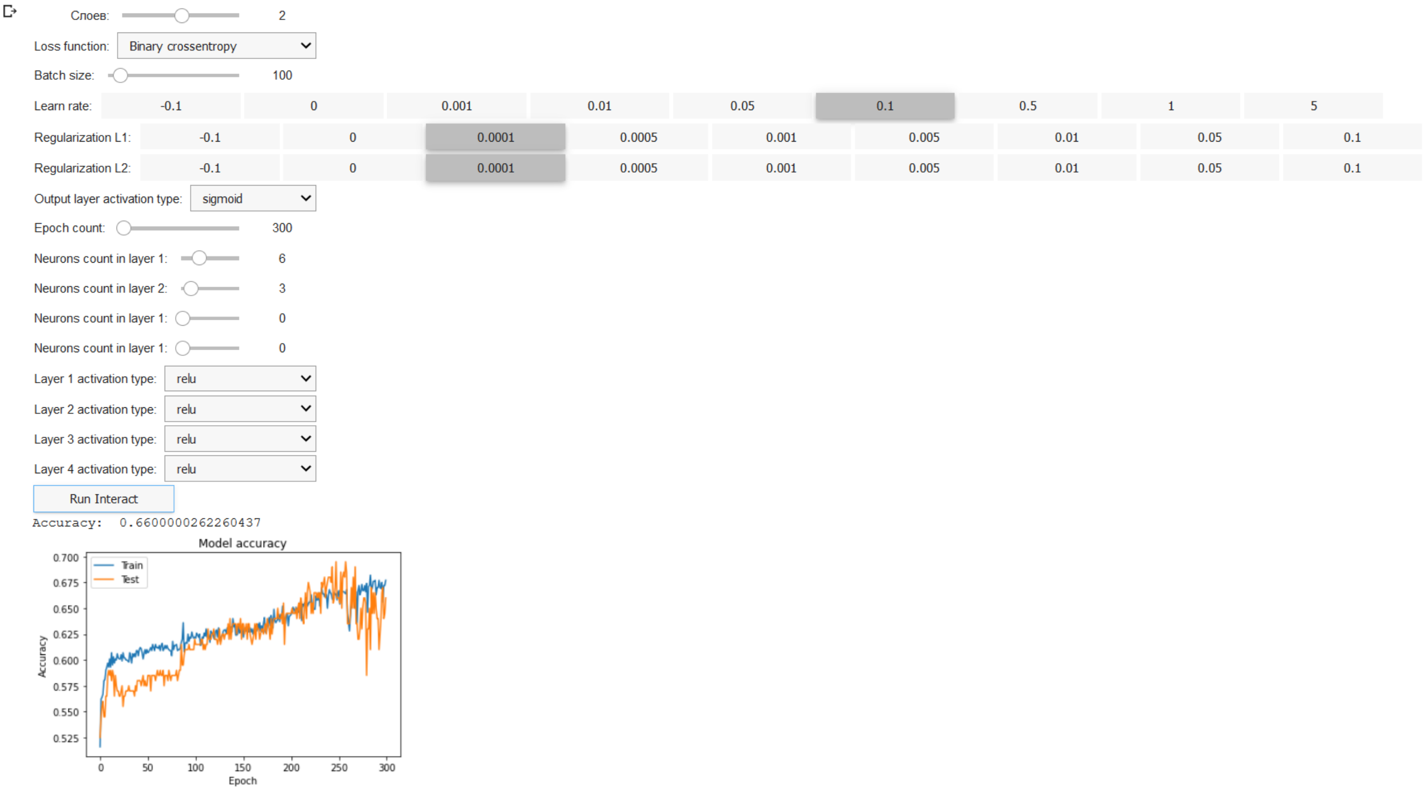
Layer 1 activation type: relu

Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.6899999976158142
```



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | [✓ 0.1] | 0.5 | 1 | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

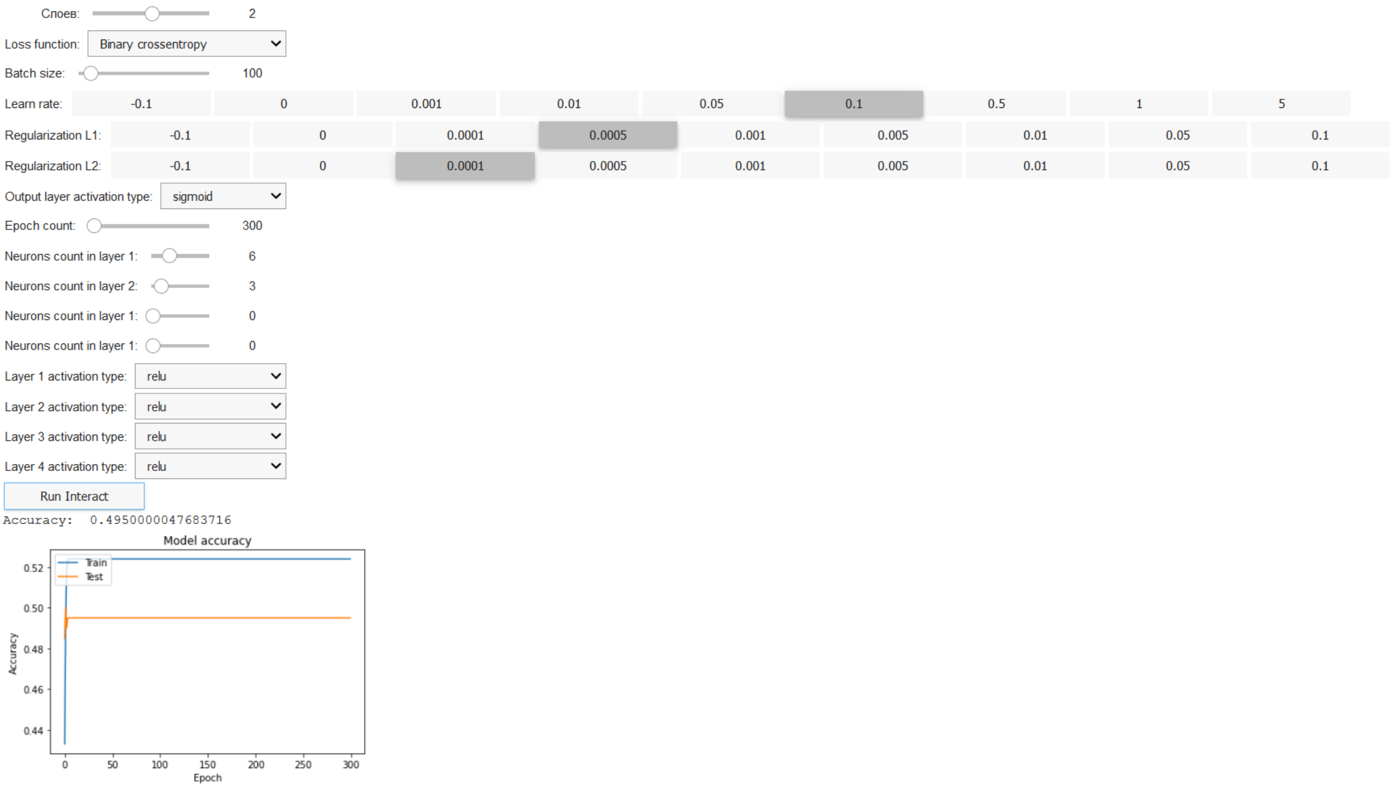
Layer 1 activation type: relu

Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.6600000262260437
```



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 [✓ 0.1] 0.5 1 5

Regularization L1: -0.1 0 0.0001 [✓ 0.0005] 0.001 0.005 0.01 0.05 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

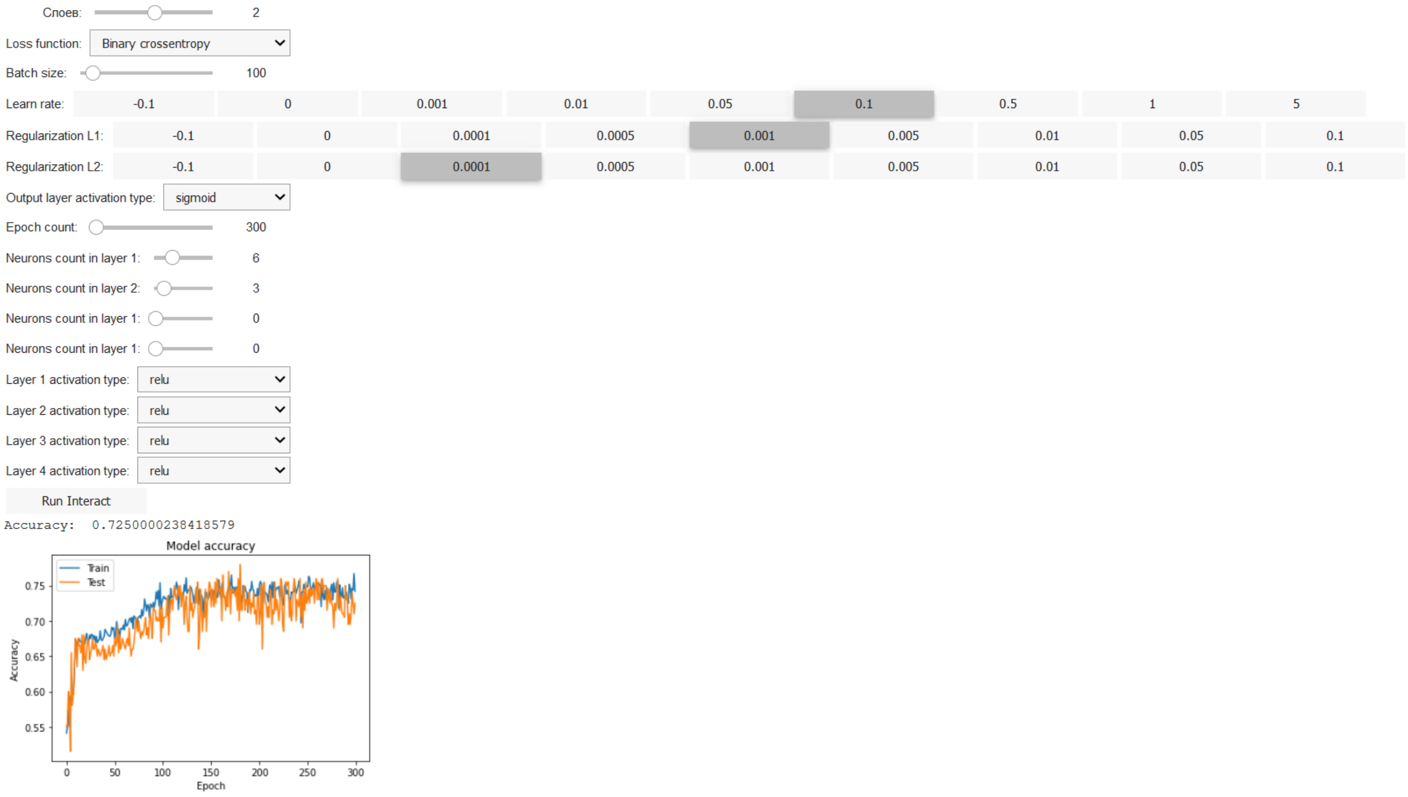
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.4950000047683716
```

Model accuracy



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 [✓ 0.1] 0.5 1 5

Regularization L1: -0.1 0 0.0001 0.0005 [✓ 0.001] 0.005 0.01 0.05 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

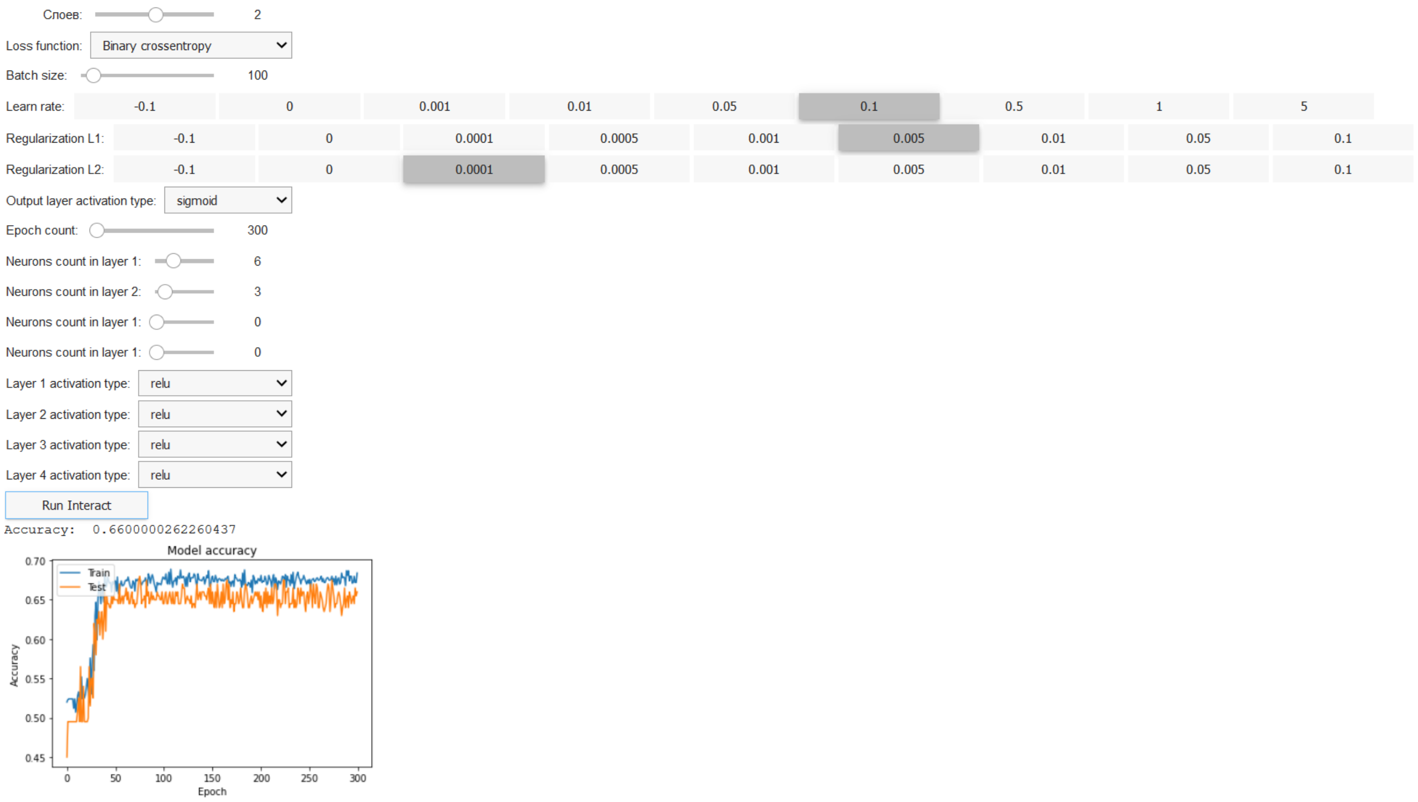
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.7250000238418579
```

Model accuracy



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 [✓ 0.1] 0.5 1 5

Regularization L1: -0.1 0 0.0001 0.0005 0.001 [✓ 0.005] 0.01 0.05 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

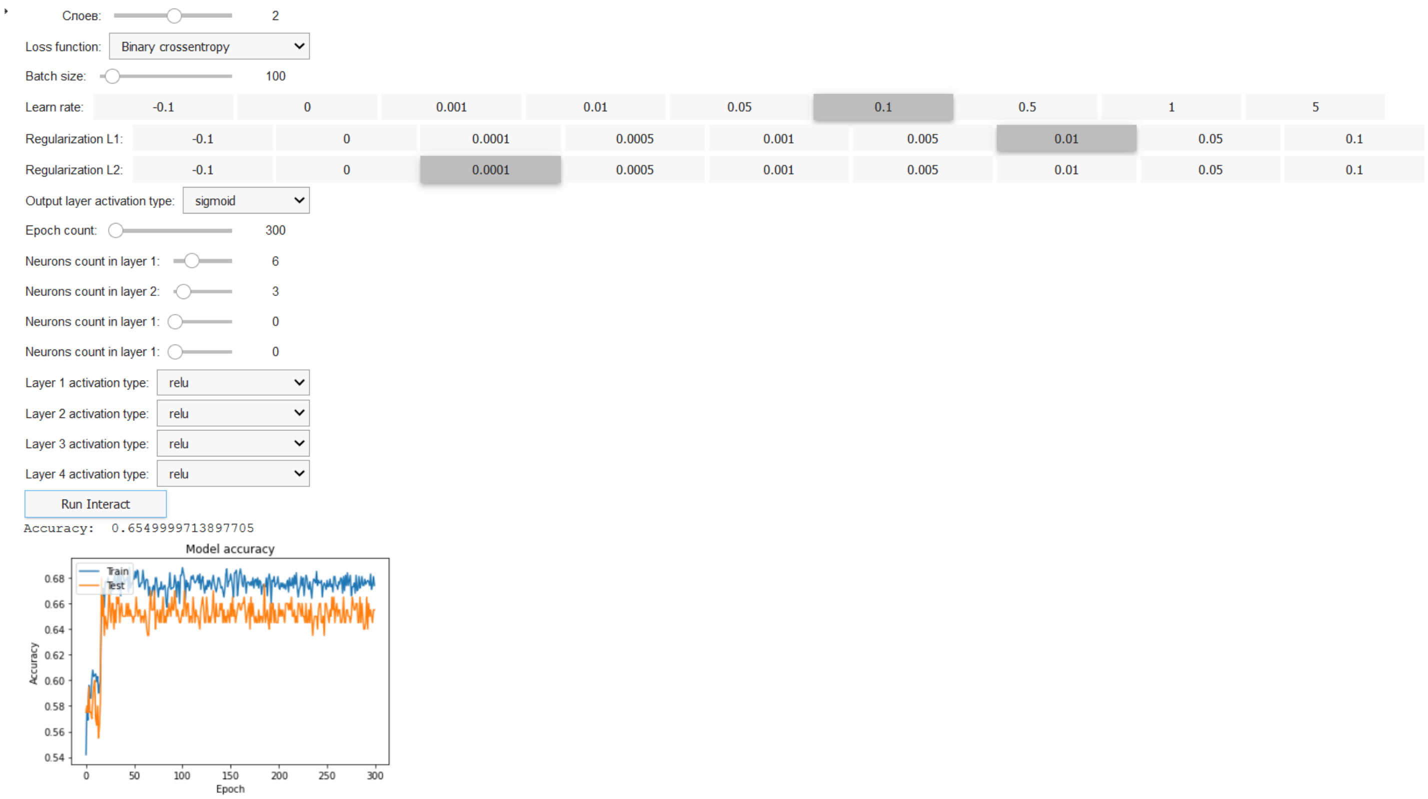
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

Run Interact

```
Accuracy:  0.6600000262260437
```



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 [✓ 0.1] 0.5 1 5

Regularization L1: -0.1 0 0.0001 0.0005 0.001 0.005 [✓ 0.01] 0.05 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

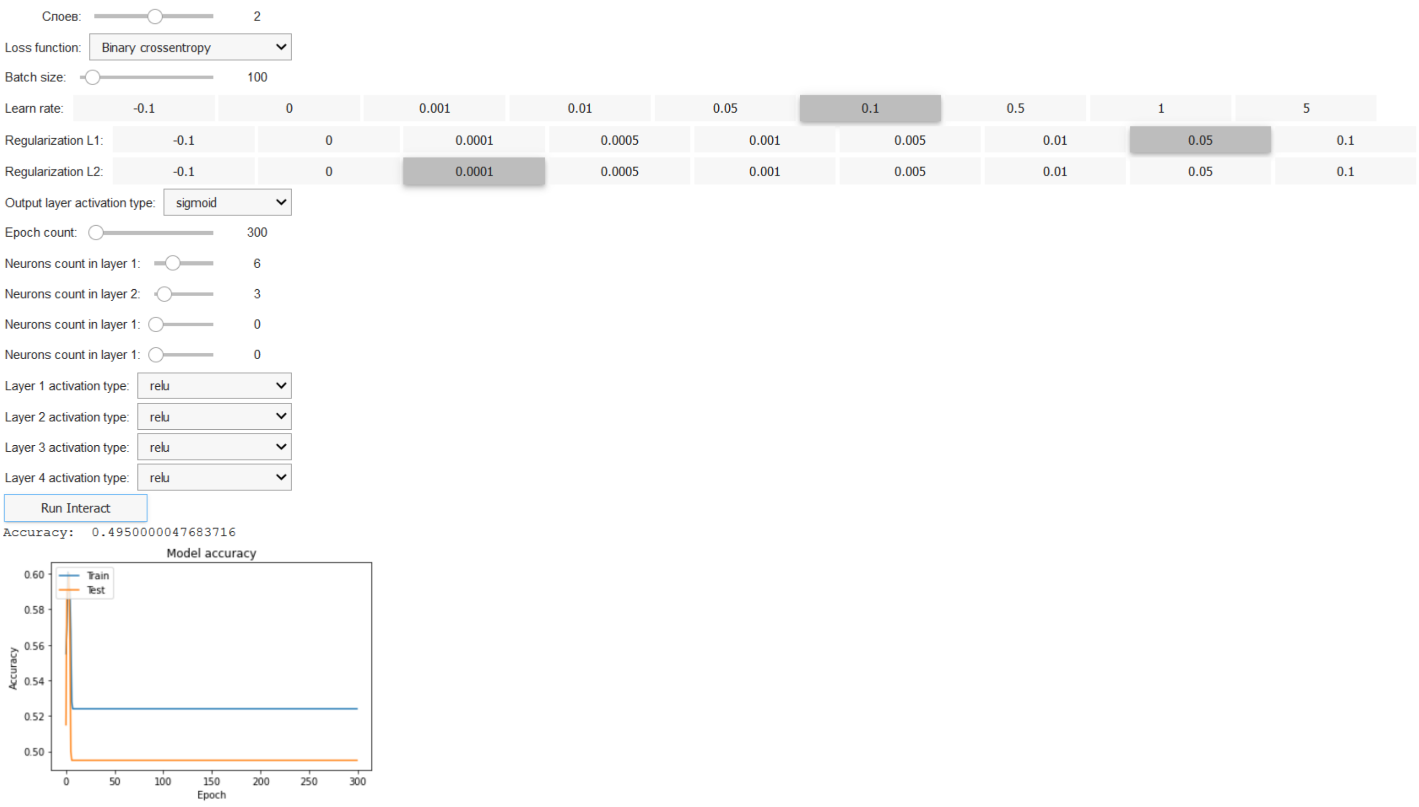
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

Run Interact

```
Accuracy:  0.6549999713897705
```



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 [✓ 0.1] 0.5 1 5

Regularization L1: -0.1 0 0.0001 0.0005 0.001 0.005 0.01 [✓ 0.05] 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

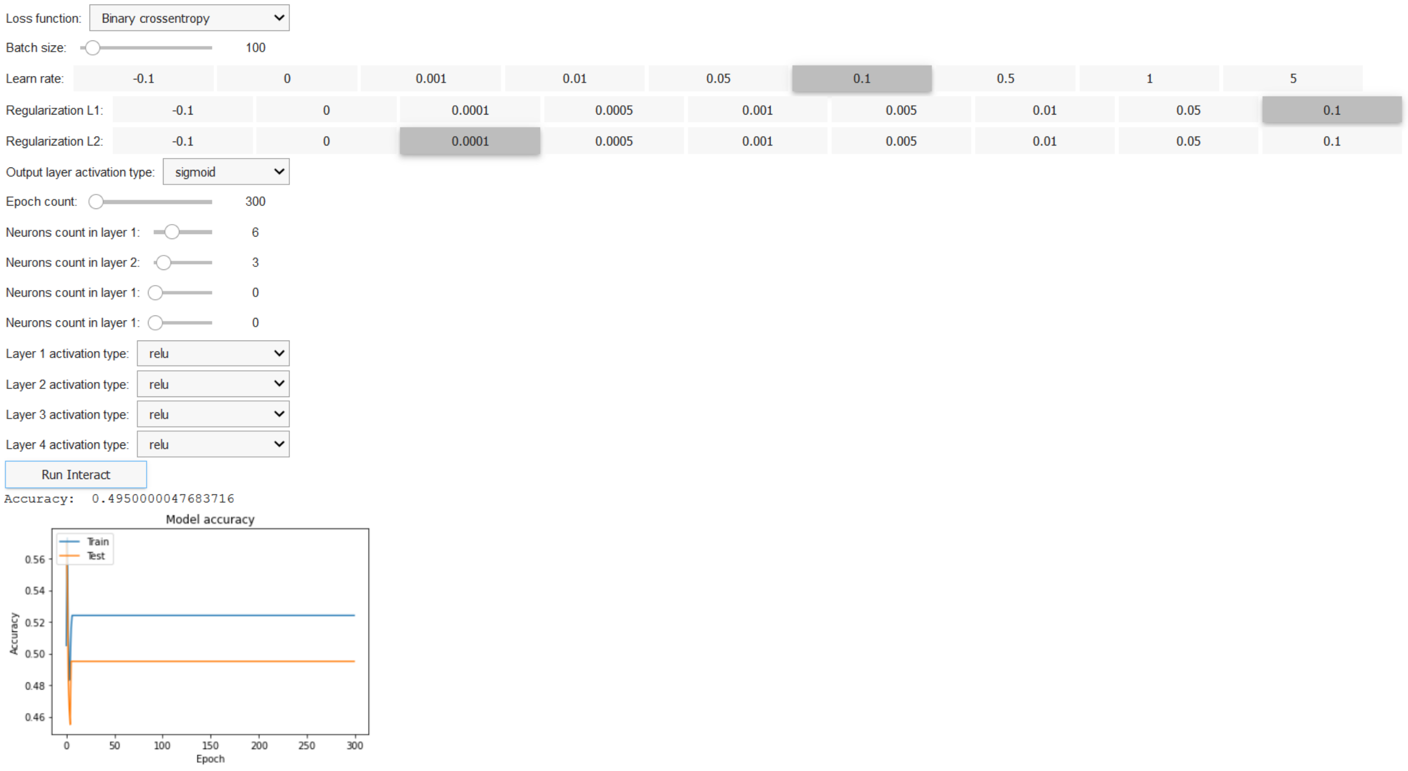
Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

Run Interact

```
Accuracy:  0.4950000047683716
```



Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 [✓ 0.1] 0.5 1 5

Regularization L1: -0.1 0 0.0001 0.0005 0.001 0.005 0.01 0.05 [✓ 0.1]

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

Layer 1 activation type: relu

Layer 2 activation type: relu

Layer 3 activation type: relu

Layer 4 activation type: relu

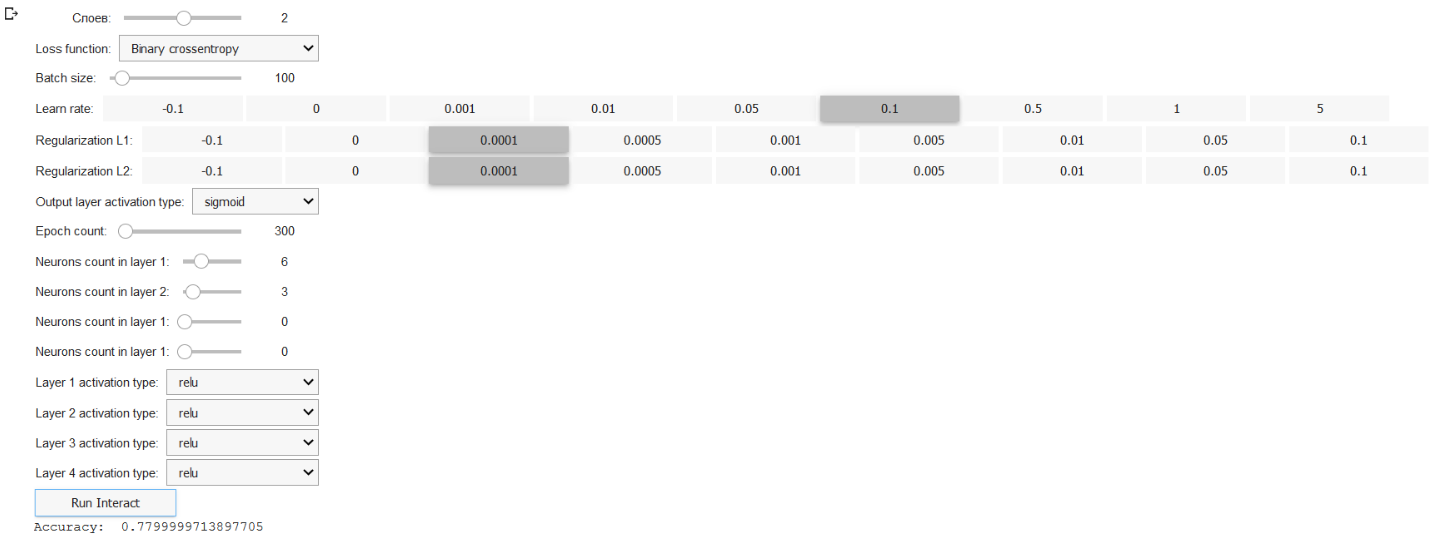
Run Interact

```
Accuracy:  0.4950000047683716
```

Highest accuracy

Learn rate: 0.1

Regularization L1: 0.001



Слоев: 2

Loss function: Binary crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | [✓ 0.1] | 0.5 | 1 | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 300

Neurons count in layer 1: 6

Neurons count in layer 2: 3

Neurons count in layer 1: 0

Neurons count in layer 1: 0

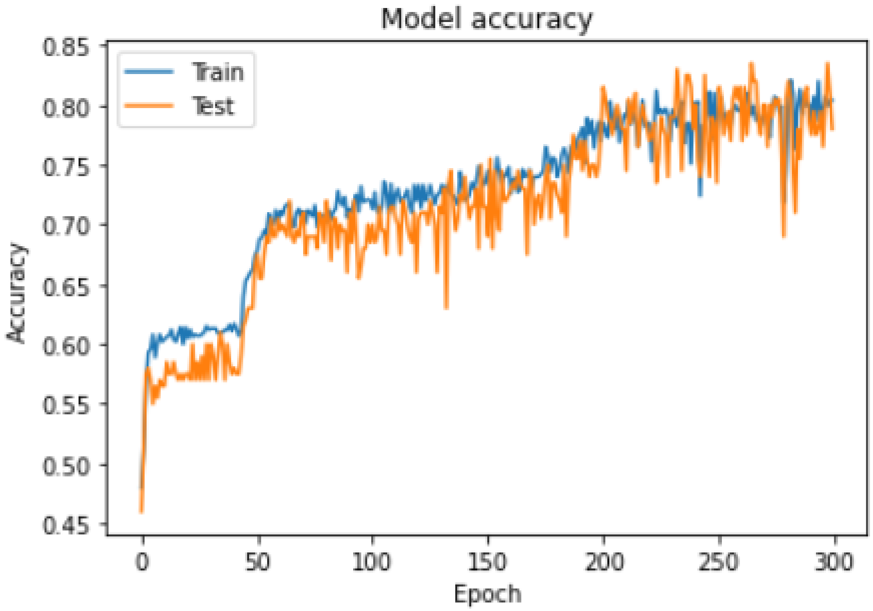
Layer 1 activation type: relu

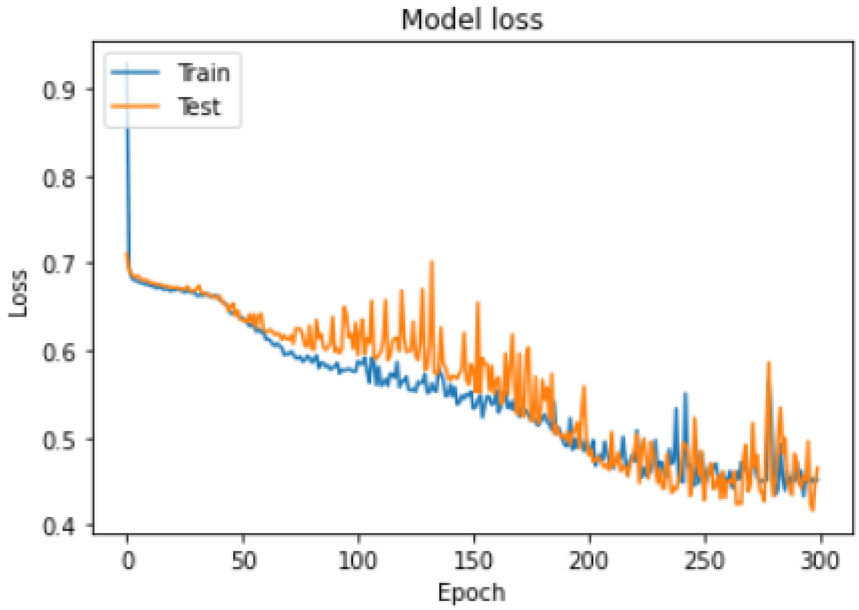
Layer 2 activation type: relu

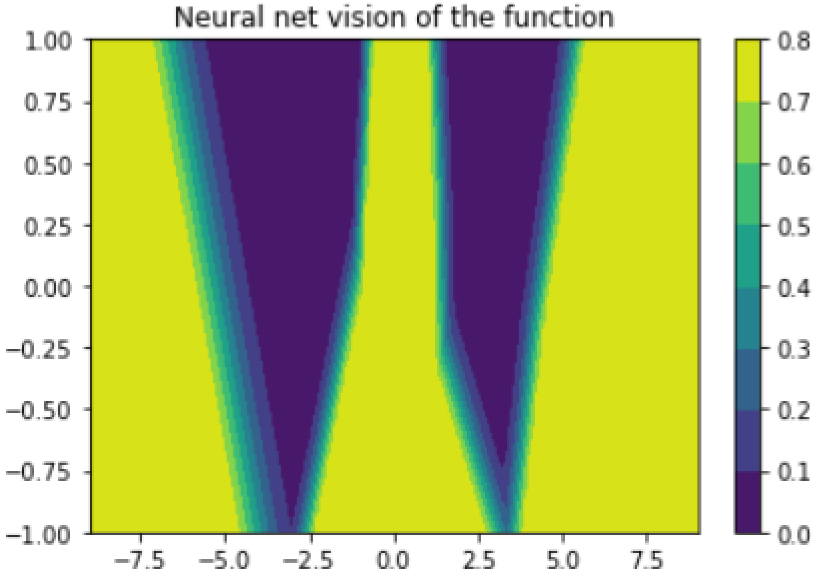
Layer 3 activation type: relu

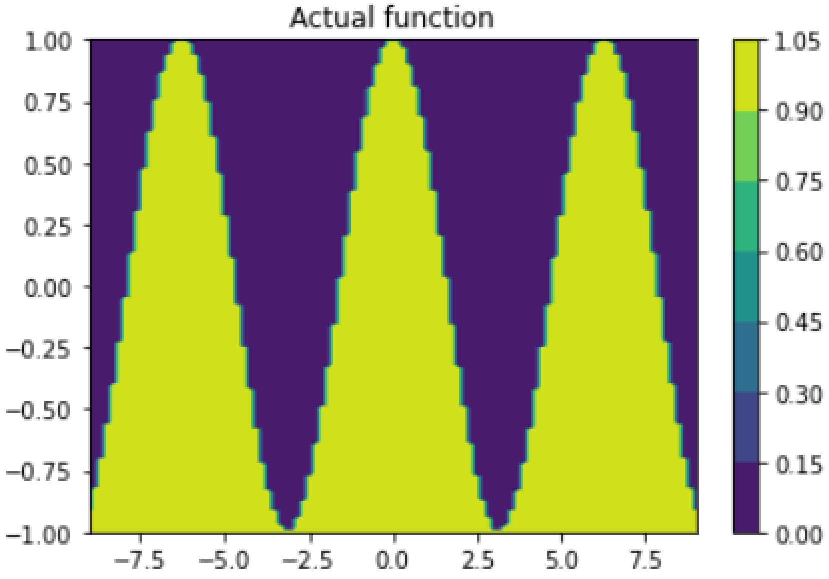
Layer 4 activation type: relu

```
Accuracy:  0.7799999713897705
```









Patterns

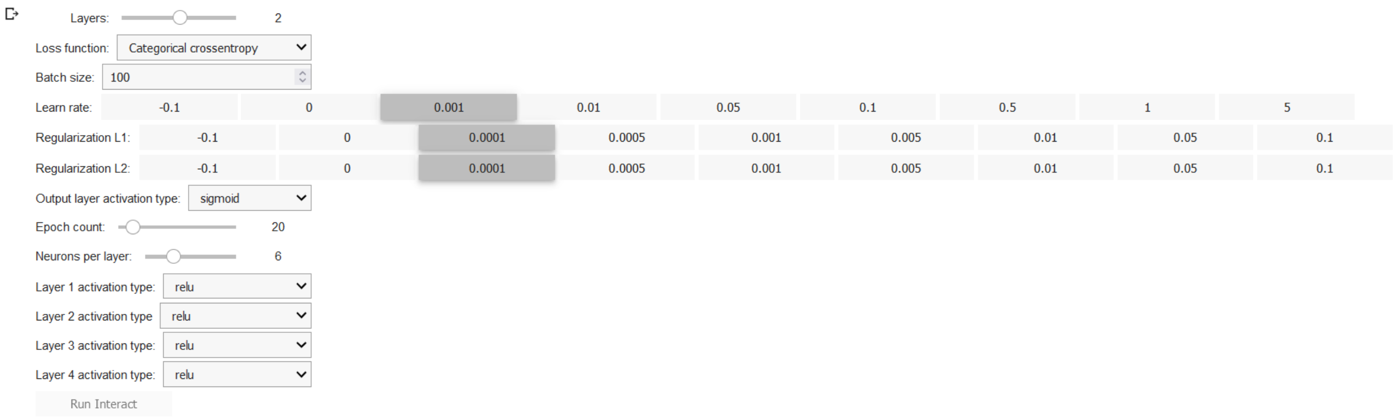
	Коэф. скорости обучения	Точность
1	-0.1	NULL
2	0	0.52
3	0.001	0.57
4	0.01	0.65
5	0.05	0.68
6	0.1	0.79
7	0.5	0.67

8	1	0.65
9	5	0.50

According to the experimental data, it can be seen that with an increase in the learning rate from 0.001 to 0.1, the accuracy increases. This contradicts the assumption made prior to the experiment.
It is also strange that the result of accuracy was obtained with a coefficient equal to zero (this should not have happened).

According to the experimental data, with an increase in the learning rate from 0.1 to 5, the accuracy decreases. This is in line with the assumption made prior to the experiment.
It is strange that the program showed some accuracy with a coefficient equal to 5, because this value is outside the range of acceptable values.

Часть 2

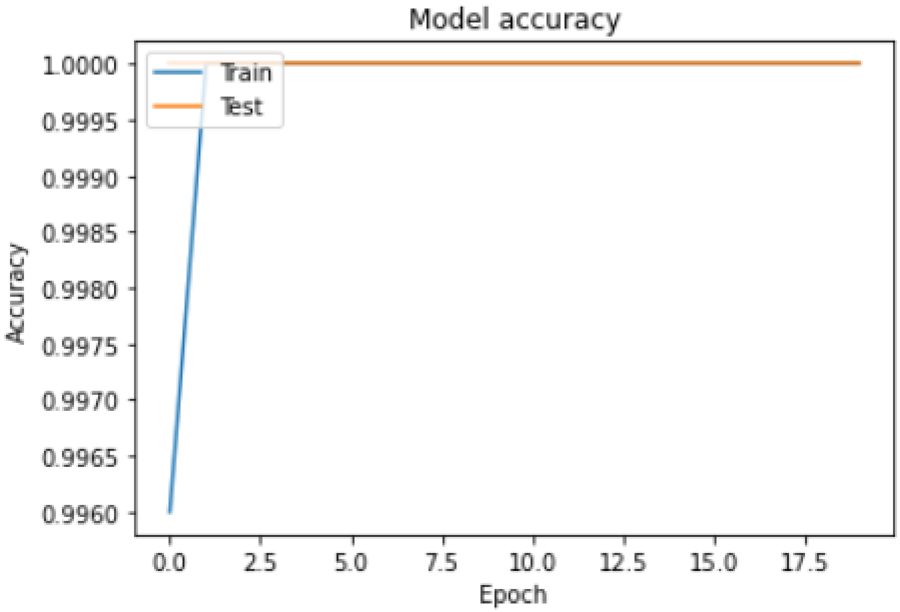


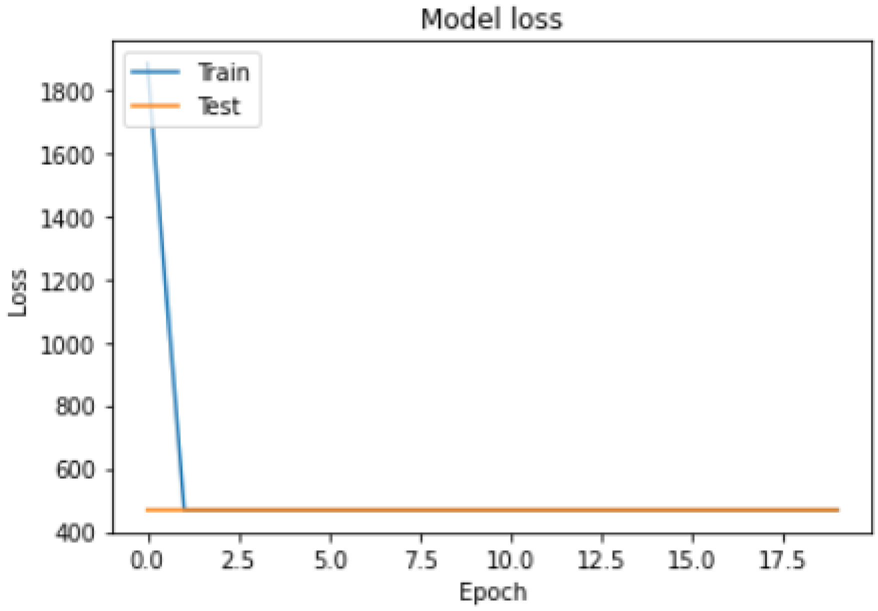
Layers: 2
Loss function: Categorical crossentropy
Batch size: 100

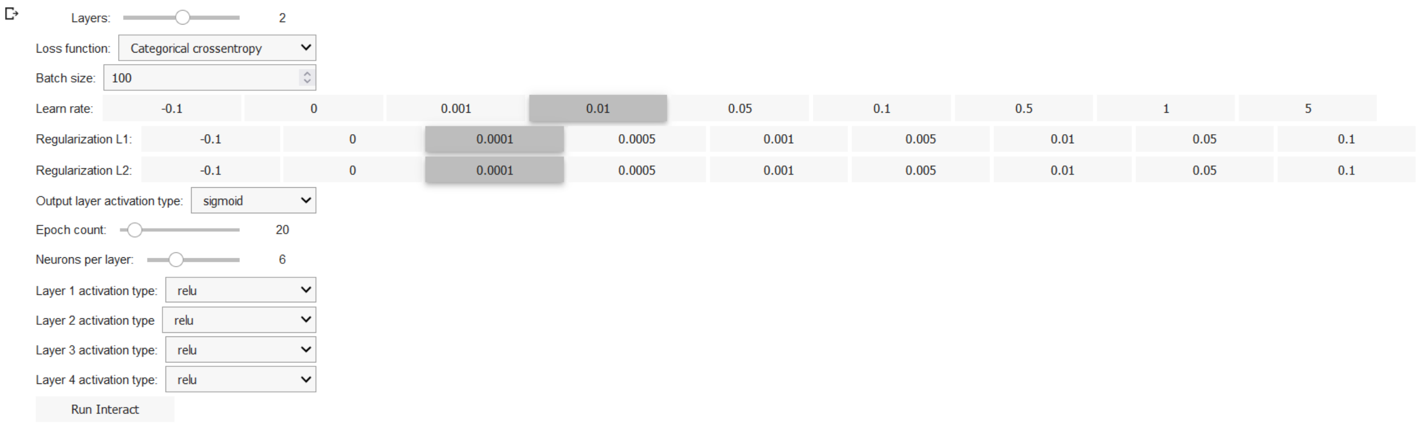
Learn rate:	-0.1	0	**0.001**	0.01	0.05	0.1	0.5	1	5
Regularization L1:	-0.1	0	**0.0001**	0.0005	0.001	0.005	0.01	0.05	0.1
Regularization L2:	-0.1	0	**0.0001**	0.0005	0.001	0.005	0.01	0.05	0.1

Output layer activation type: sigmoid
Epoch count: 20
Neurons per layer: 6
Layer 1 activation type: relu
Layer 2 activation type relu
Layer 3 activation type: relu
Layer 4 activation type: relu

```
Accuracy:  1.0
```







Layers: 2

Loss function: Categorical crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 [✓ 0.01] 0.05 0.1 0.5 1 5

Regularization L1: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 20

Neurons per layer: 6

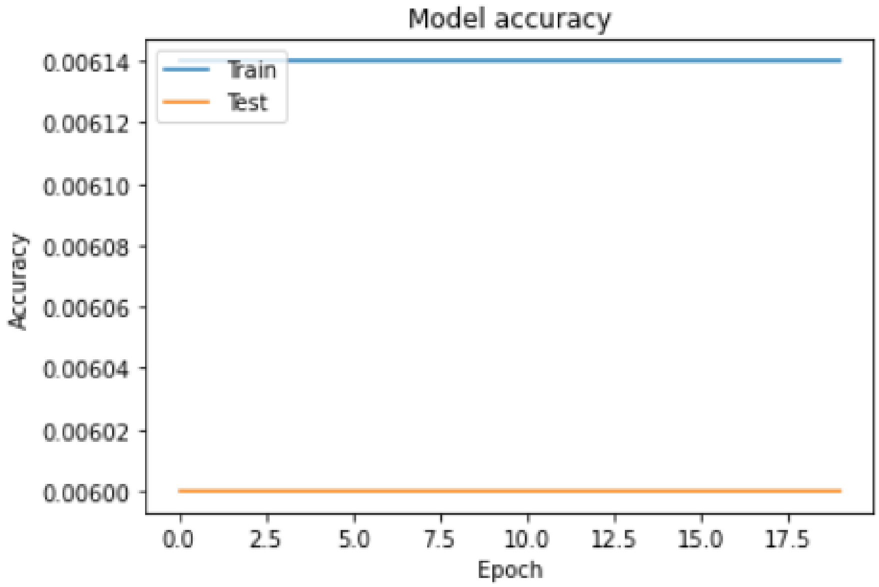
Layer 1 activation type: relu

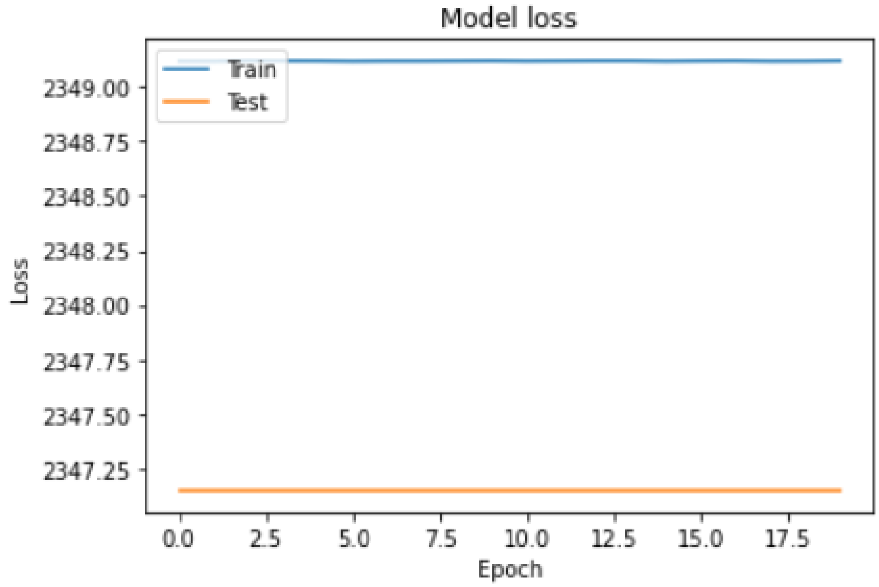
Layer 2 activation type relu

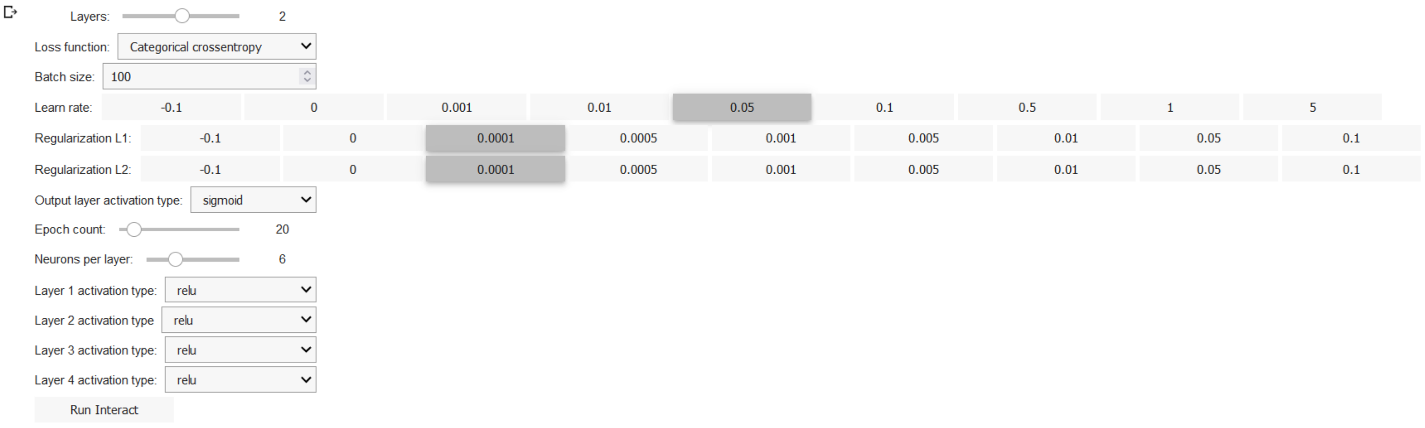
Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  0.006000000052154064
```







Layers: 2

Loss function: Categorical crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | [✓ 0.05] | 0.1 | 0.5 | 1 | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 20

Neurons per layer: 6

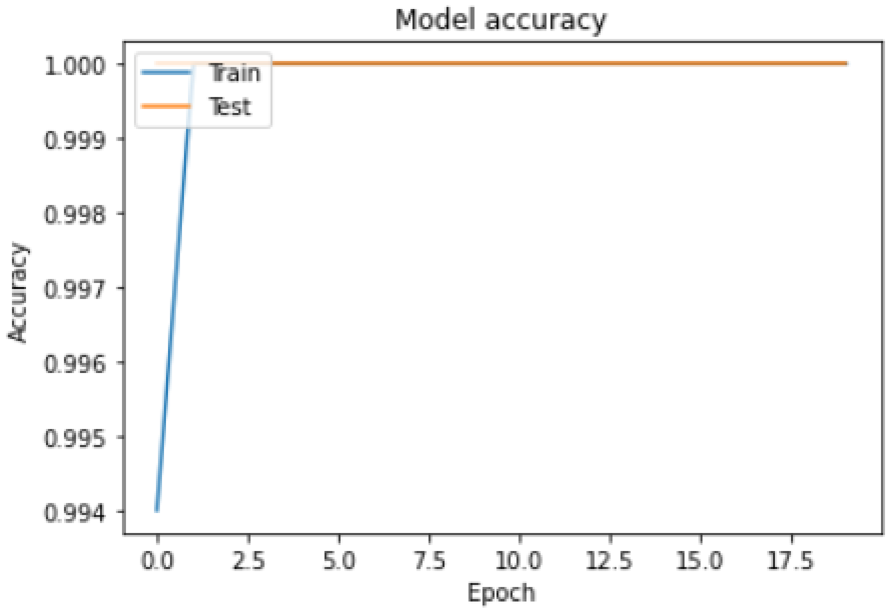
Layer 1 activation type: relu

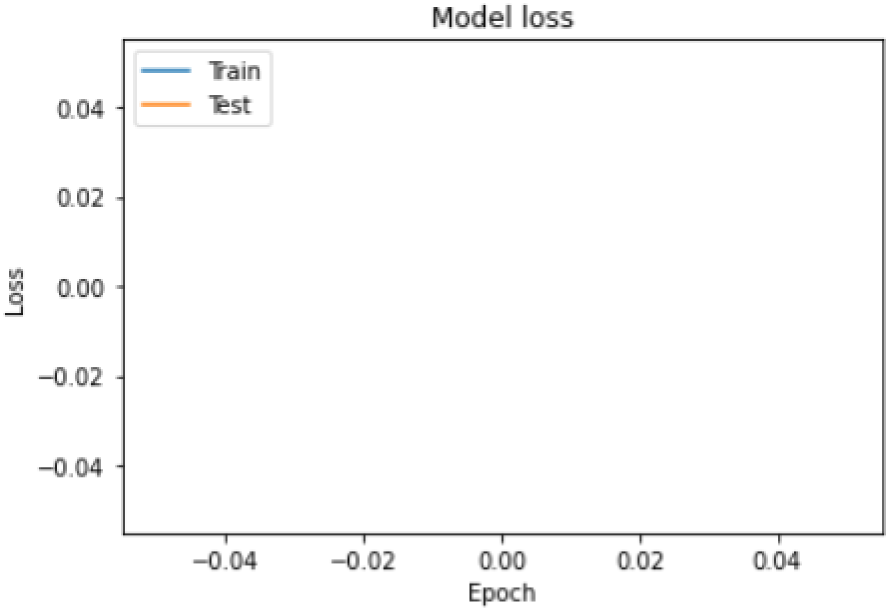
Layer 2 activation type relu

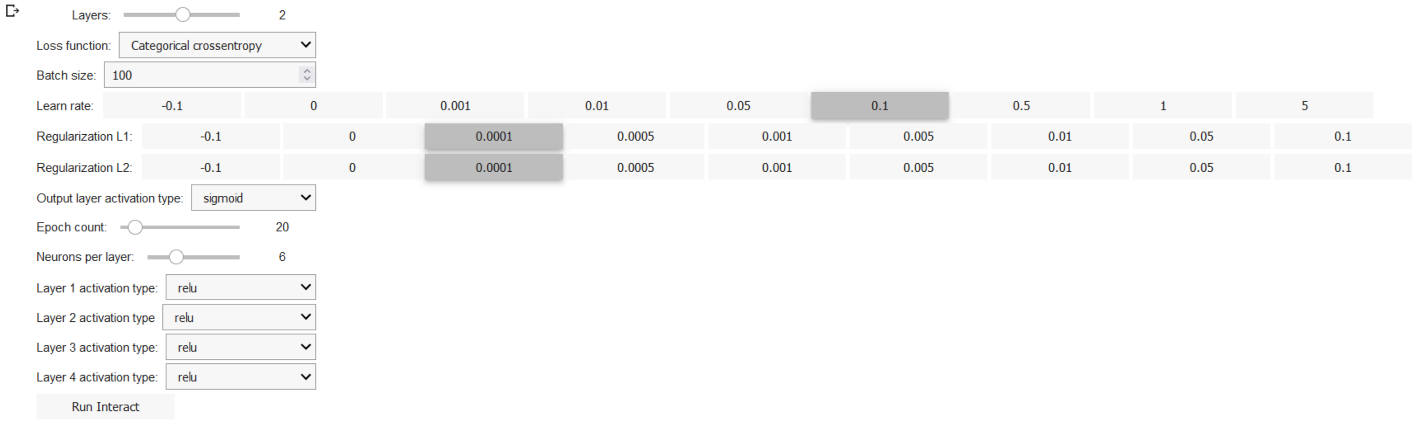
Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  1.0
```


Model accuracy


Model loss



Layers: 2

Loss function: Categorical crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 [✓ 0.1] 0.5 1 5

Regularization L1: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 20

Neurons per layer: 6

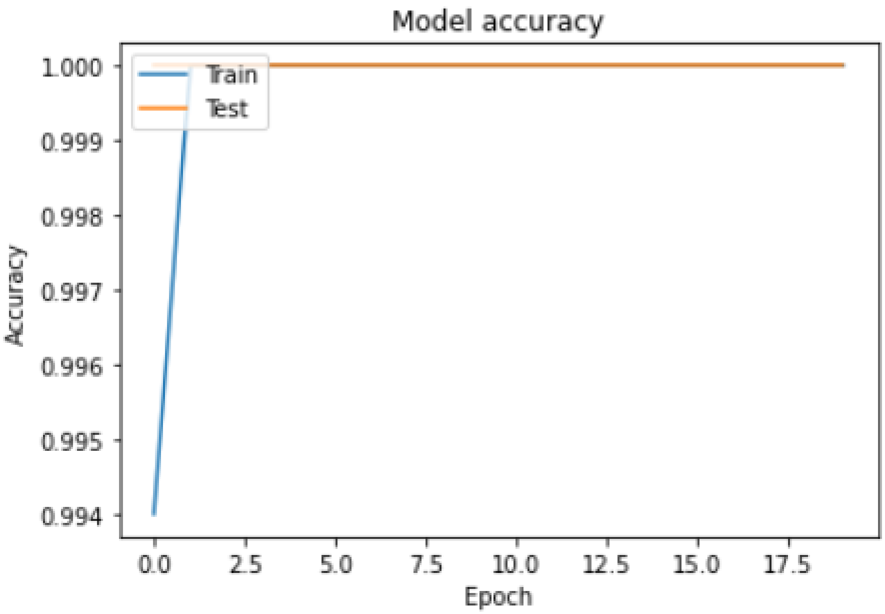
Layer 1 activation type: relu

Layer 2 activation type relu

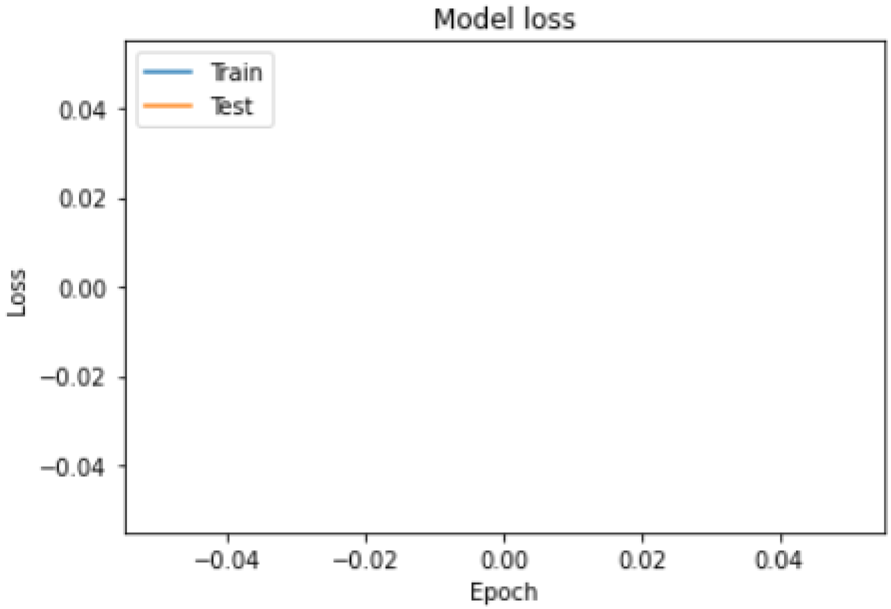
Layer 3 activation type: relu

Layer 4 activation type: relu

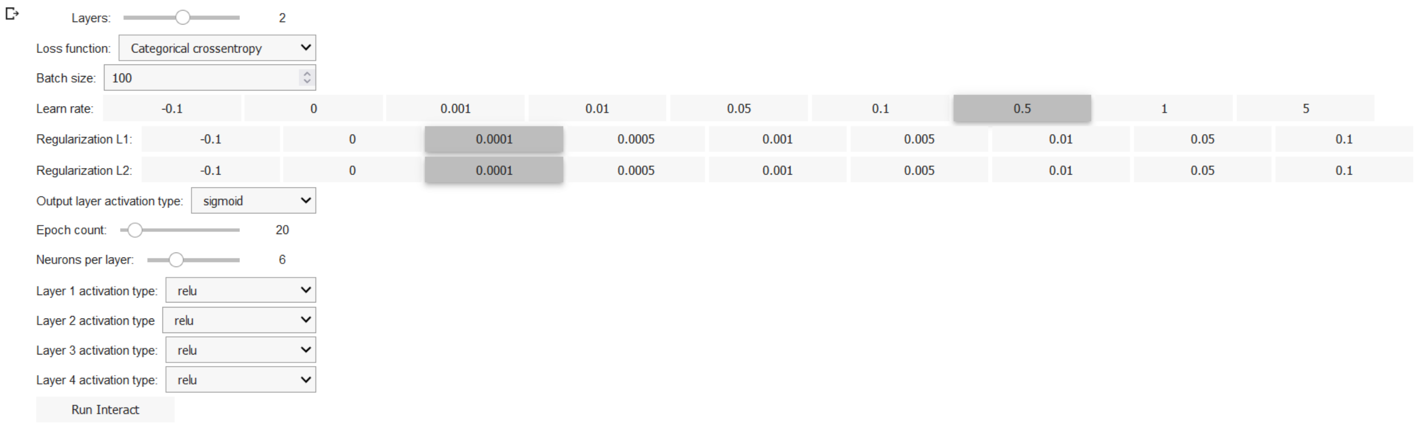
```
Accuracy:  1.0
```



Model accuracy



Model loss



Layers: 2

Loss function: Categorical crossentropy

Batch size: 100

Learn rate: -0.1 0 0.001 0.01 0.05 0.1 [✓ 0.5] 1 5

Regularization L1: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Regularization L2: -0.1 0 [✓ 0.0001] 0.0005 0.001 0.005 0.01 0.05 0.1

Output layer activation type: sigmoid

Epoch count: 20

Neurons per layer: 6

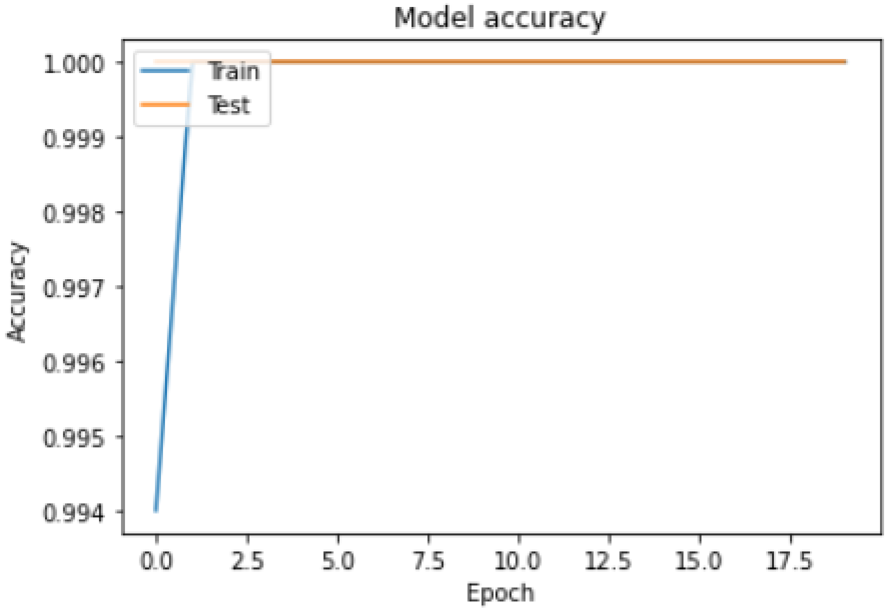
Layer 1 activation type: relu

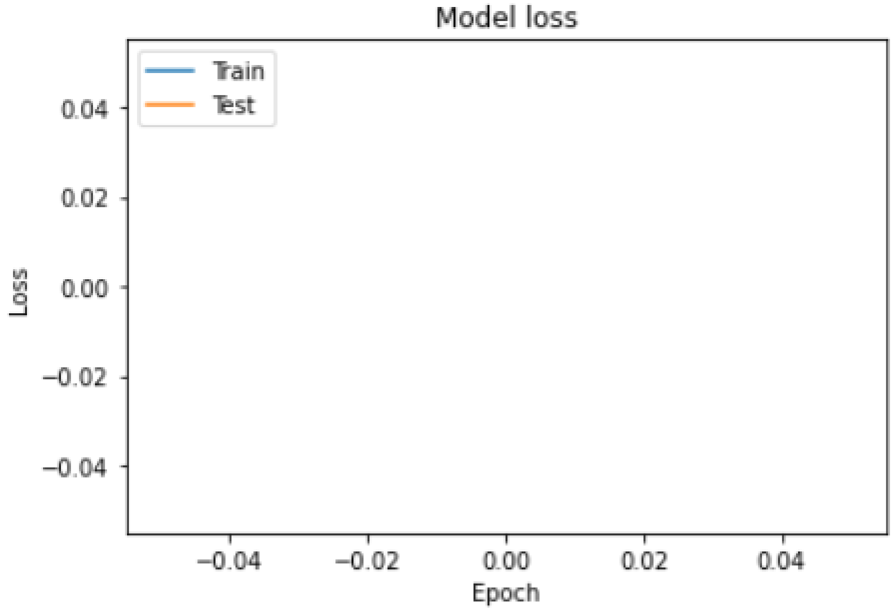
Layer 2 activation type: relu

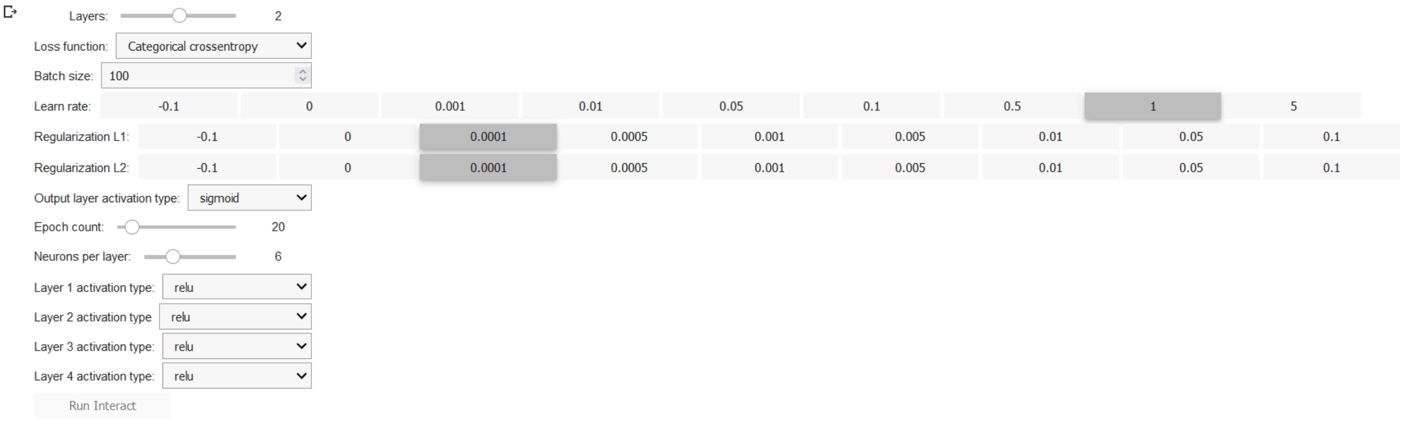
Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  1.0
```







Layers: 2

Loss function: Categorical crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | 0.1 | 0.5 | [✓ 1] | 5

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 20

Neurons per layer: 6

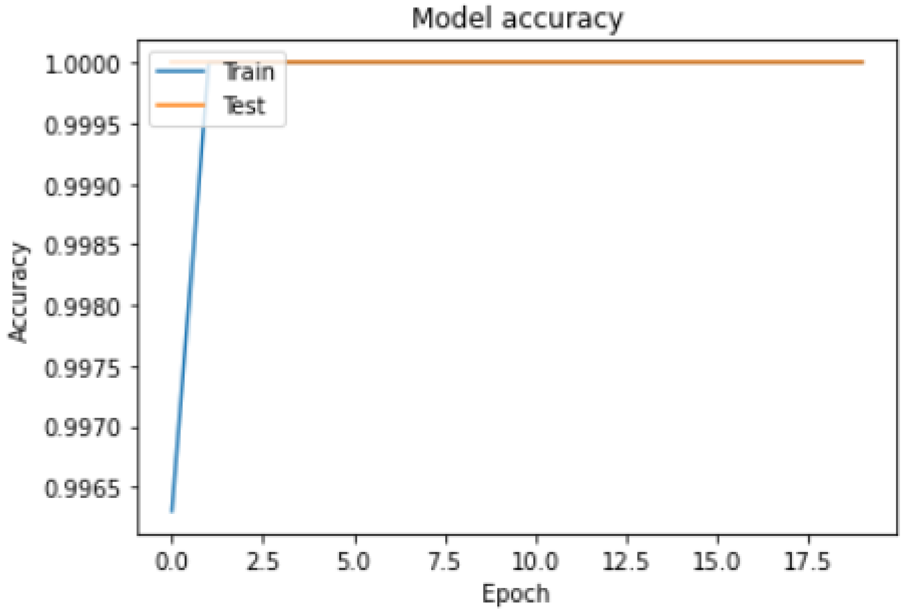
Layer 1 activation type: relu

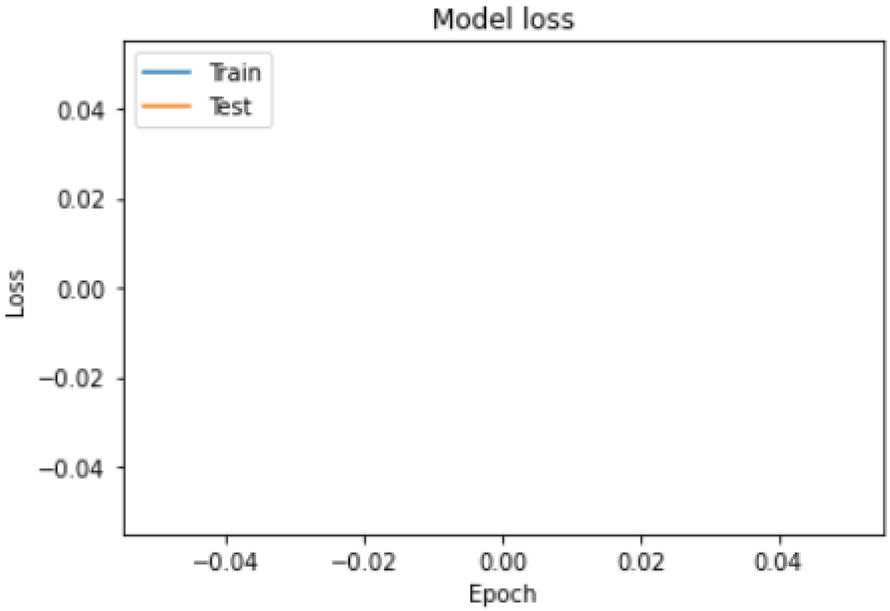
Layer 2 activation type relu

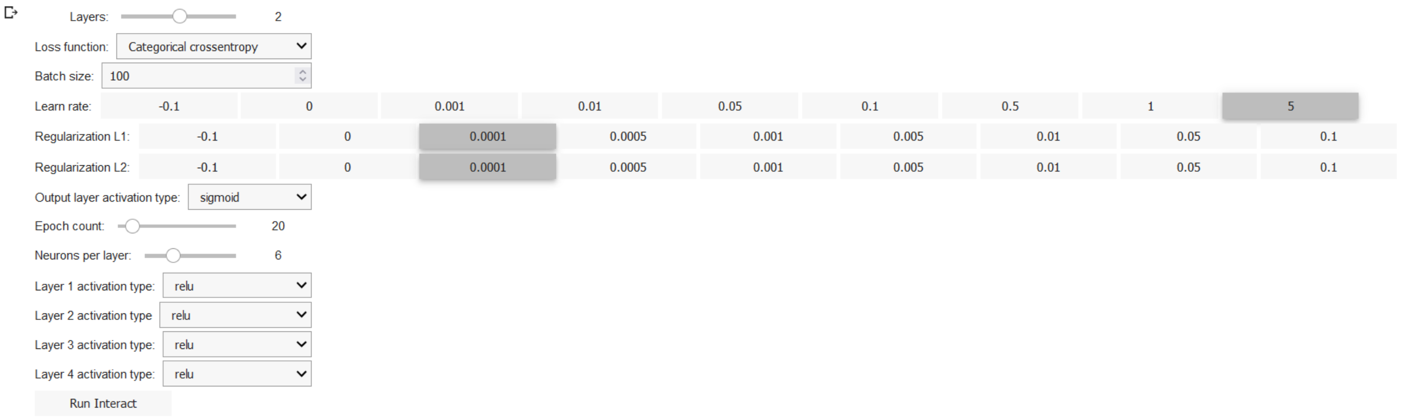
Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  1.0
```







Layers: 2

Loss function: Categorical crossentropy

Batch size: 100

Learn rate: -0.1 | 0 | 0.001 | 0.01 | 0.05 | 0.1 | 0.5 | 1 | [✓ 5]

Regularization L1: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Regularization L2: -0.1 | 0 | [✓ 0.0001] | 0.0005 | 0.001 | 0.005 | 0.01 | 0.05 | 0.1

Output layer activation type: sigmoid

Epoch count: 20

Neurons per layer: 6

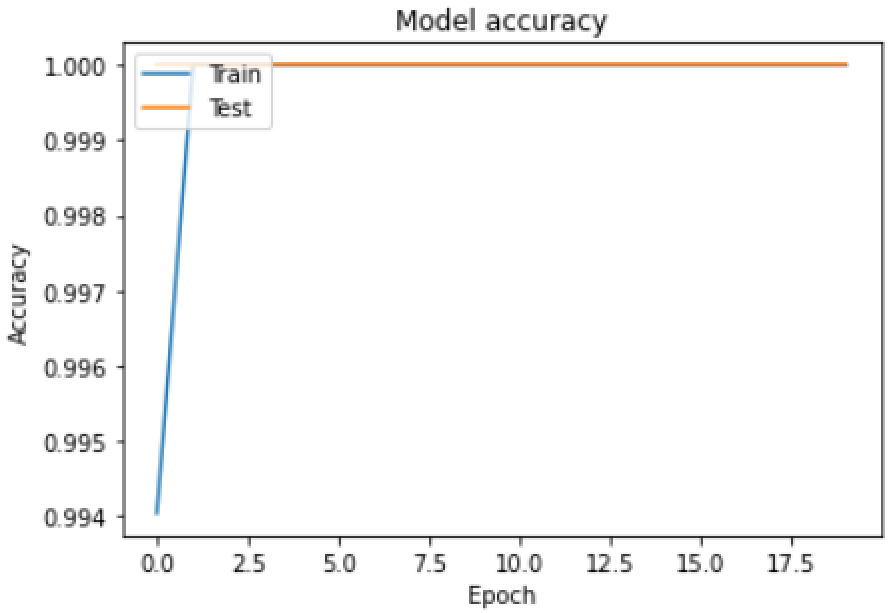
Layer 1 activation type: relu

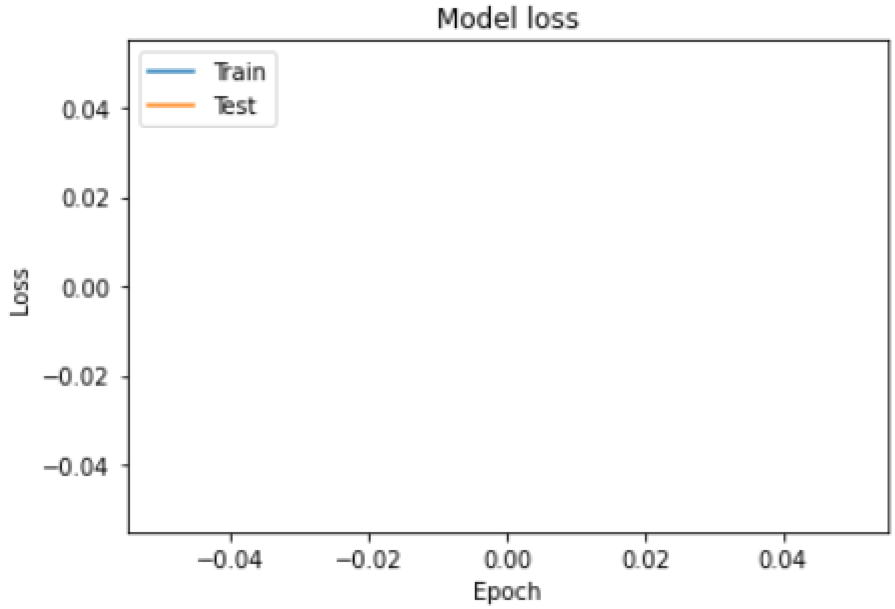
Layer 2 activation type relu

Layer 3 activation type: relu

Layer 4 activation type: relu

```
Accuracy:  1.0
```





Patterns

Regardless of the L1 regularization, the accuracy during all experiments was equal to 1.

Frankly, I don’t know if it really should be this way.

We considered L1 regularization only as an idea. I have not seen anywhere that the “L1” coefficient is used somewhere in the formula, so it never became clear to me how changing its value should affect the training accuracy.

Вывод

The learning rate coefficient is the amount of weight correction at each iteration. As it increases, the speed of the algorithm increases, but the accuracy decreases.

L1 regularization is one of the ways to combat overfitting by adding some penalty for large values of the coefficients. This additionally leads to the selection of the most important features by bringing the coefficients at the least important to zero.