

## Подбор гиперпараметров

Рассматривалась следующая задача:

$$\begin{aligned}iq_t + q_{xx} + |q|^2 q(1 - \alpha|q|^2 + \beta|q|^4) &= 0, \quad x \in [x_L, x_R], \quad t \in [0, t_{max}], \\q(x, 0) &= q_0(x), \\q(x_L, t) &= q(x_R, t), \\q_x(x_L, t) &= q_x(x_R, t),\end{aligned}\tag{1}$$

имеющая решение в виде светлого солитона:

$$\begin{aligned}q(x, t) &= \sqrt{\frac{\mu e^{(x-2kt-x_0)\sqrt{\mu}}}{(\frac{1}{2}e^{(x-2kt-x_0)\sqrt{\mu}} + 1)^2 - \frac{\alpha_0\mu}{3}e^{2(x-2kt-x_0)\sqrt{\mu}}}} e^{i(kx-\omega t+\theta_0)}, \\ \mu &= 4(k^2 - \omega).\end{aligned}\tag{2}$$

В данной части работы осуществлялся подбор гиперпараметров модели.

Определялись оптимальные значения для количества точек коллокации, точек для начальных и граничных условий, скорости обучения, числа слоёв, числа нейронов, определялся также вид функции активации.

На каждой итерации случайным образом выбиралось новое сочетание гиперпараметров из заданного пространства поиска:

- **num\_domain** выбиралось случайным образом из равномерного целочисленного распределения в диапазоне от 10 000 до 30 000;
- **num\_boundary** выбиралось случайным образом из равномерного целочисленного распределения в диапазоне от 1 до 1 000;
- **num\_initial** выбиралось случайным образом из равномерного целочисленного распределения в диапазоне от 1 до 5 000;
- **learning\_rate** выбиралось случайным образом из логарифмически равномерного распределения в диапазоне от  $1 \times 10^{-4}$  до  $5 \times 10^{-2}$ ;
- **num\_layers** выбиралось случайным образом из равномерного целочисленного распределения в диапазоне от 1 до 4;
- **num\_nodes** выбиралось случайным образом из равномерного целочисленного распределения в диапазоне от 5 до 100;
- **activation** выбиралось случайным образом из категориального распределения с возможными значениями: **sin**, **sigmoid** и **tanh**.

Затем эти параметры использовались для обучения модели нейронной сети, и ошибка оценивалась на тестовом наборе данных.

Область определения осталась прежней:  $x \in [-25, 25]$ . В качестве параметров для отладки были взяты:  $\alpha = 0$ ,  $\beta = 0$ ,  $t_{max} = 1$ ,  $\alpha_0 = 0$ ,  $k = 1$ ,  $\omega = 0.88$ ,  $x_0 = 0$ ,  $\theta_0 = 0$ . Обучение проводилось на протяжении 10000 итераций, использовался оптимизатор Adam.

Входные и выходные данные не подвергались дополнительной обработке.

Были получены следующие результаты:

num_domain	num_boundary	num_initial	learning_rate	num_layers	num_nodes	activation
19953	818	3061	0.0121	4	19	sin

Таблица 1: Подобранные гиперпараметры

	$Lw_{1\_max}$	$Lw_{1\_mean}$	$Lw_{2\_max}$	$Lw_{2\_mean}$	$Rel_h$
Старые гиперпараметры	2.3954	1.8832	3.2085	2.3782	0.0408
Новые гиперпараметры	0.0754	0.0376	0.1124	0.0588	0.0039

Таблица 2: Прирост точности при применении новых гиперпараметров