

Федеральное государственное автономное образовательное
учреждение высшего образования

Университет ИТМО

Системы искусственного интеллекта

Лабораторная работа 4

Вариант 2

Выполнил: Щербаков Александр Валерьевич,
студент группы Р33151

Преподаватель: Королёва Юлия Александровна

2023

Цель

Изучение гиперпараметров нейронной сети, понимание влияния различных гиперпараметров на точность.

Задание

Часть 1 представляет собой распознавание основных математических функций с иллюстрациями нейросетевого видения функций для обучения.

Часть 2 представляет собой распознавание простых изображений.

Вариант: 2 – (312192)

| Var | Part1 func | Part2 data | Hyperparameters |
|-----|--------------------------|------------|-------------------------------|
| 2 | Cos(x) X: -9..9 Y: -1..1 | CIFAR100 | Learn rate, regularization L1 |

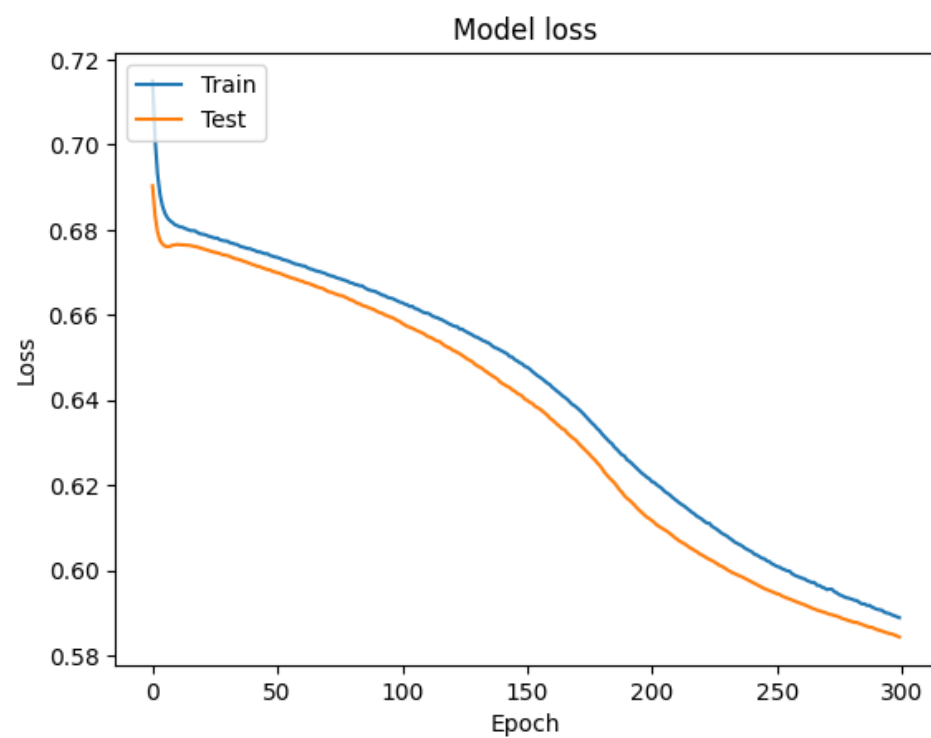
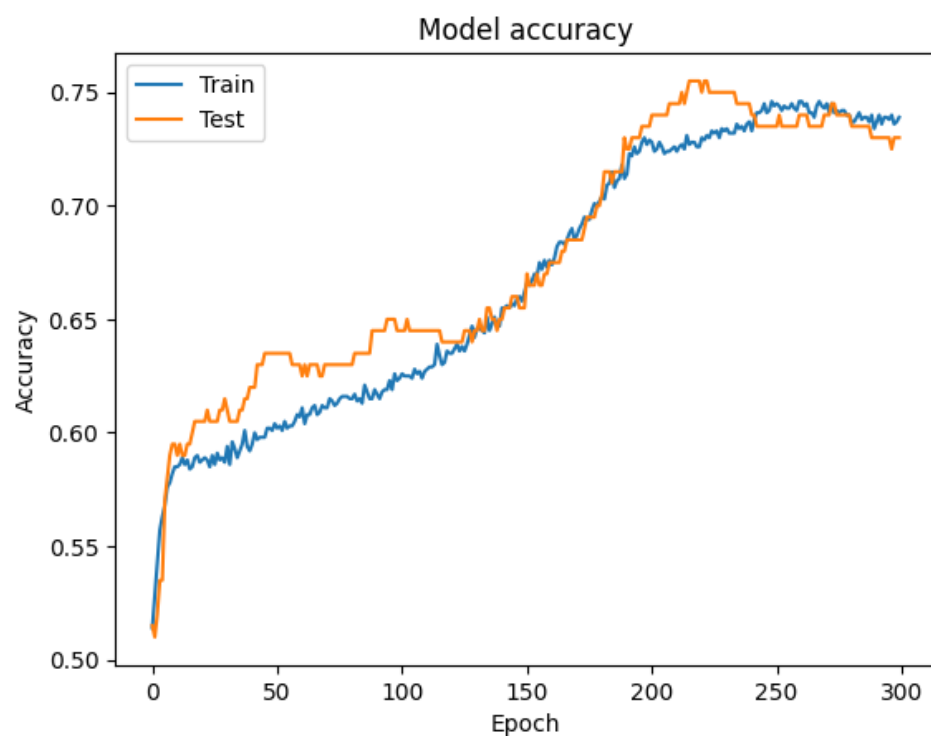
Описание гиперпараметров

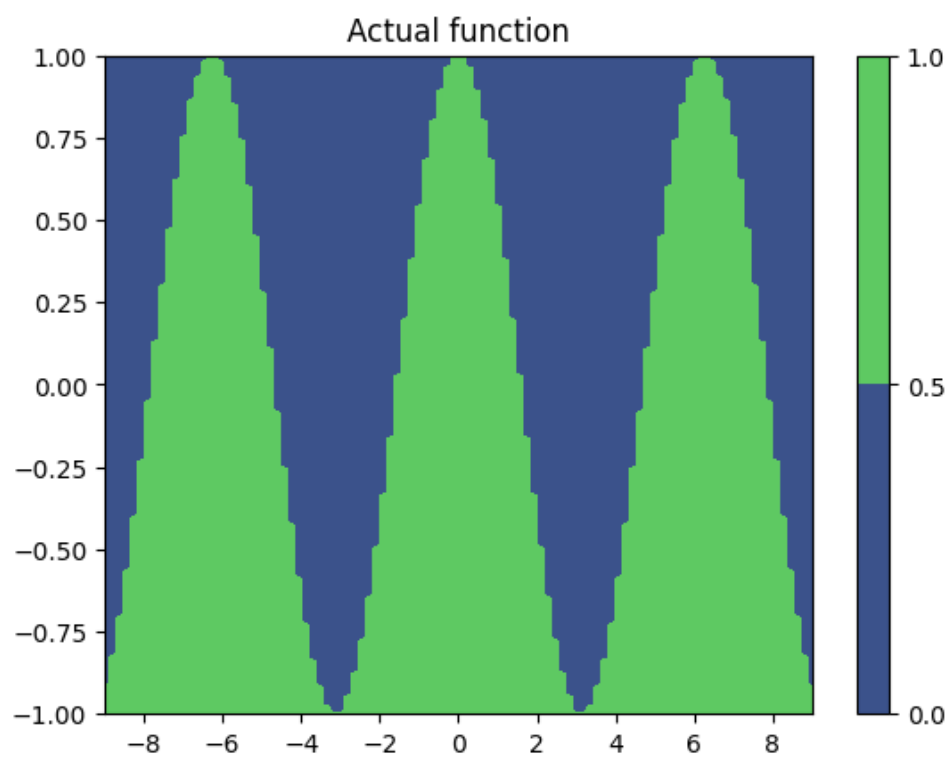
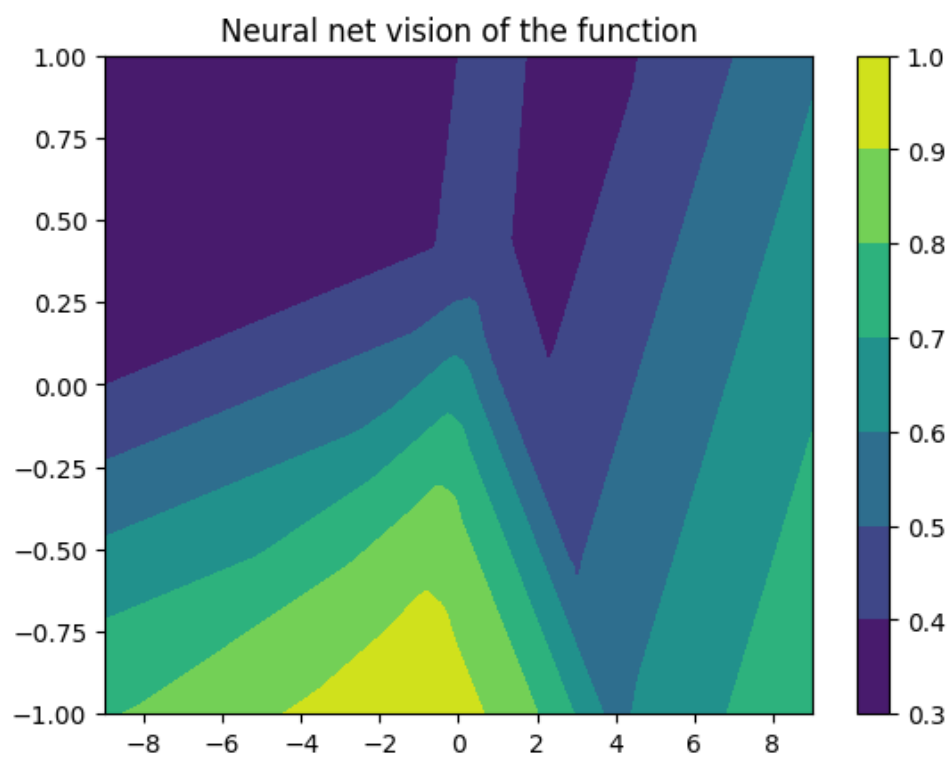
В работе представлены такие гиперпараметры, как:

- 1) Layer count – количество слоев в нейросети, в контексте сети с **SGD**
- 2) Neurons count per layer (actually it's not hyperparameter but structure parameter) – количество нейронов на слой.
- 3) Learn rate - гиперпараметр, определяющий порядок того, как мы будем корректировать наши веса с учётом функции потерь в градиентном спуске. Чем ниже величина, тем медленнее мы движемся по наклонной.
- 4) Regularization L1 and L2 – параметры регуляризации. Формально, регуляризация добавляет в функцию потерь дополнительное слагаемое налагающее штраф за сложность модели.
- 5) Output layer activation type – тип активации выходных нейронов.
- 6) Layer activation type – тип активации нейронов. Формально, функция активации.
- 7) Loss function type – тип функции потерь.
- 8) Epoch count – количество проходов для обучения.

Результат

В первой части работы удалось добиться точности порядка 75%, с признаками небольшого переобучения при параметрах $LR=0.001$ $L1=0.0005$





Во второй части работы удалось добиться значения точности порядка 92% при **Learn rate** = 0.01, а **L1** = 0.0001

