

Отчет по лабораторной работе № 2

«Применение многослойной нейронной сети для аппроксимации функций»

студента Шамаева Сергея группы Б21-514. Дата сдачи: _____

Ведущий преподаватель: _____ оценка: _____ подпись: _____

Вариант № _____

Цель работы: изучение математической модели многослойной нейронной сети и решение с её помощью задачи аппроксимации функций.

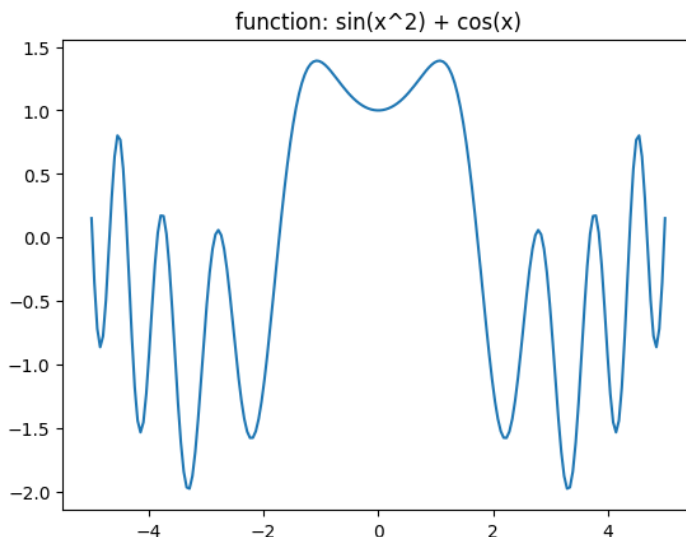
1. Подготовка данных

Аппроксимируемая функция	Число входов	Число выходов	Диапазон изменения аргументов
$\sin(x^2) + \cos(x)$	1	1	$[-5; 5]$

Формирование обучающей, валидационной и тестовой выборок:

	Обучающая	Валидационная	Тестовая	Всего
%	60	30	10	100
Объём выборки	120	60	20	200

График аппроксимируемой функции:



Предобработка данных:

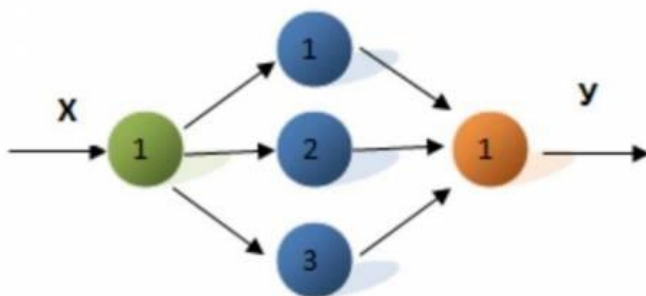
	Метод	Параметры метода	Формула расчёта
Предобработка входов	-		
Предобработка выходов	-		

2. Обучение и тестирование нейронной сети с одним скрытым слоем

Параметры архитектуры сети:

Число входов	Число выходов	Число нейронов в скрытом слое	Функция активации нейронов скрытого слоя	Функция активации выходного нейрона
1	1	40	$y = \tanh x$	Linear $y = h$

Схема нейронной сети:



Параметры обучения:

Метод обучения	Скорость обучения α	Режим обучения	Функция потерь
GD	0.04	stochastic	Quadratic loss

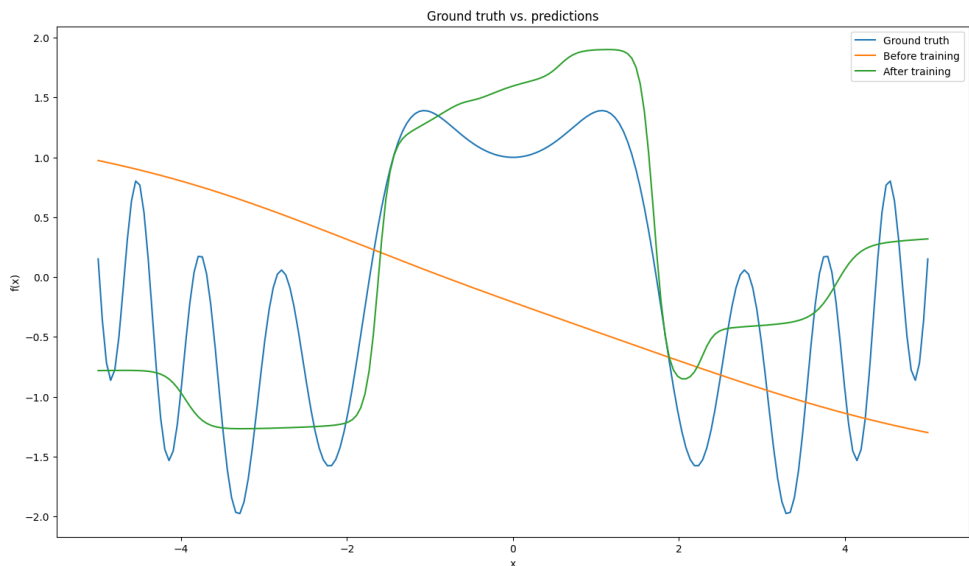
Метод инициализации сети: инициализация Хавьера

Критерий обучения:

$$E(w) = \frac{1}{n} \sum_{i=1}^n E^{(i)}(w) \rightarrow \min_w$$

Критерий останова: количество эпох.

Зависимость выхода $y(x)$ сети от входа сети (изобразить три графика: до обучения, после обучения и график аппроксимируемой функции):

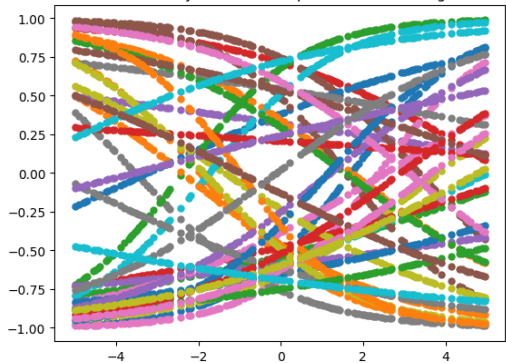


Зависимость выходов $y_k(x)$ нейронов скрытого слоя от входа сети (изобразить на одном графике):

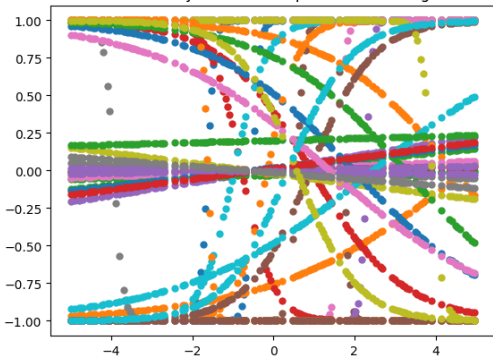
До обучения

После обучения

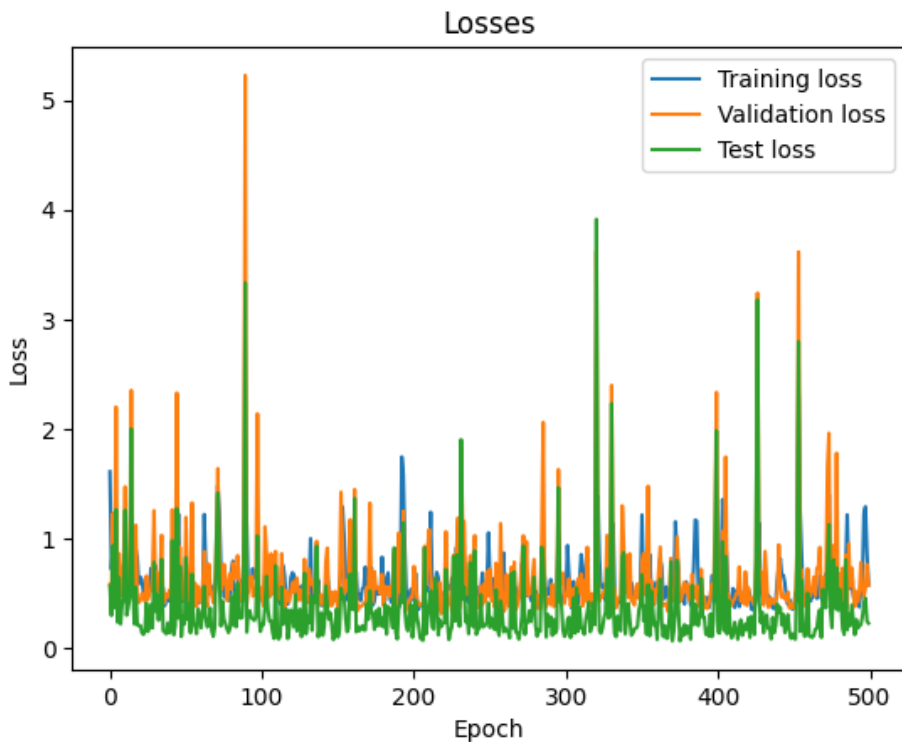
Hidden layer neuron outputs before training



Hidden layer neuron outputs after training

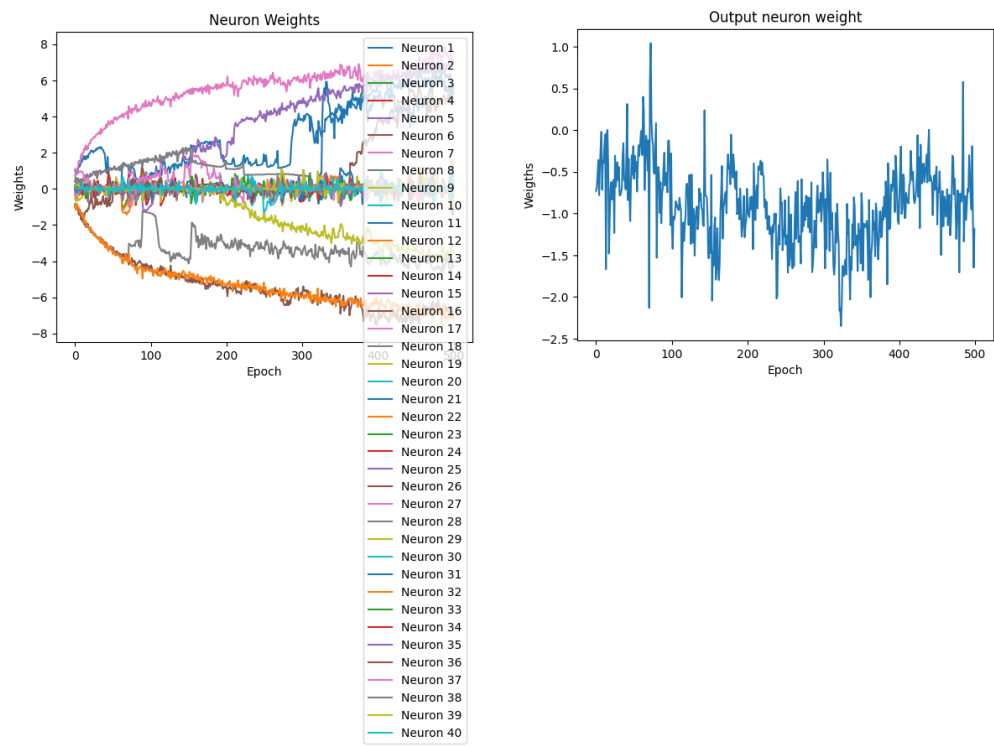


Зависимость ошибки сети $E(\tau)$ на обучающей, валидационной и тестовой выборках от времени обучения:



Отметить на графике начало переобучения (если наблюдается)

Зависимость синаптических коэффициентов сети $w(\tau)$ от времени обучения:



Показатели качества обученной нейросетевой модели:

	Обучающая	Валидационная	Тестовая
Макс. абс. ошибка	0.579	0.609	0.377
С.к.о. ошибки	0.498	0.569	0.223
RMSE	0.706	0.754	0.472

Обученная нейросетевая модель *обладает* способностью к генерализации данных. Для улучшения качества аппроксимации требуется использовать *сеть с меньшим числом нейронов, изменить параметры метода обучения, изменить режим обучения.*

3. Улучшение качества аппроксимации

Параметры архитектуры сети:

Число входов	Число выходов	Число нейронов в скрытом слое	Функция активации нейронов скрытого слоя	Функция активации выходного нейрона
1	1	10	$y = \tanh x$	Linear $y = h$

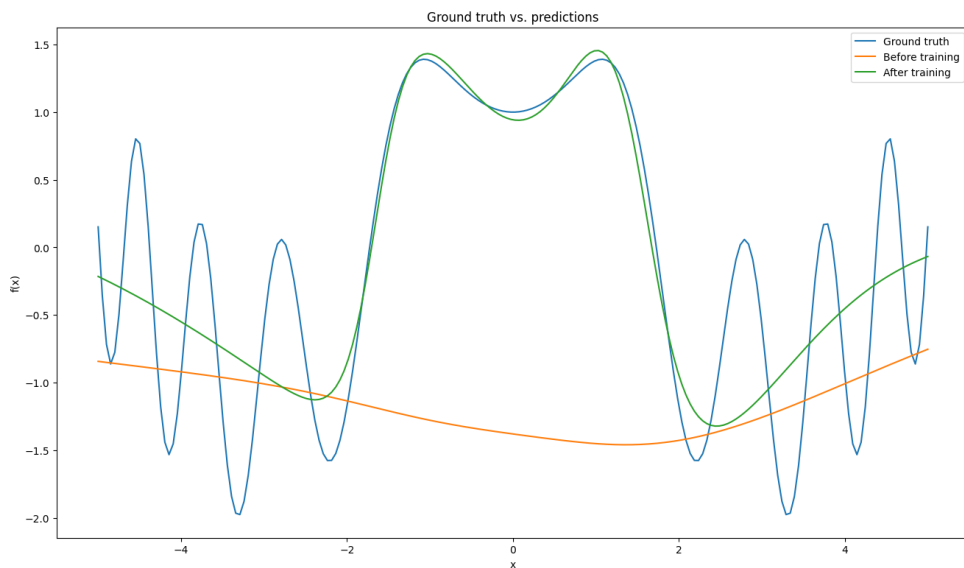
Параметры обучения:

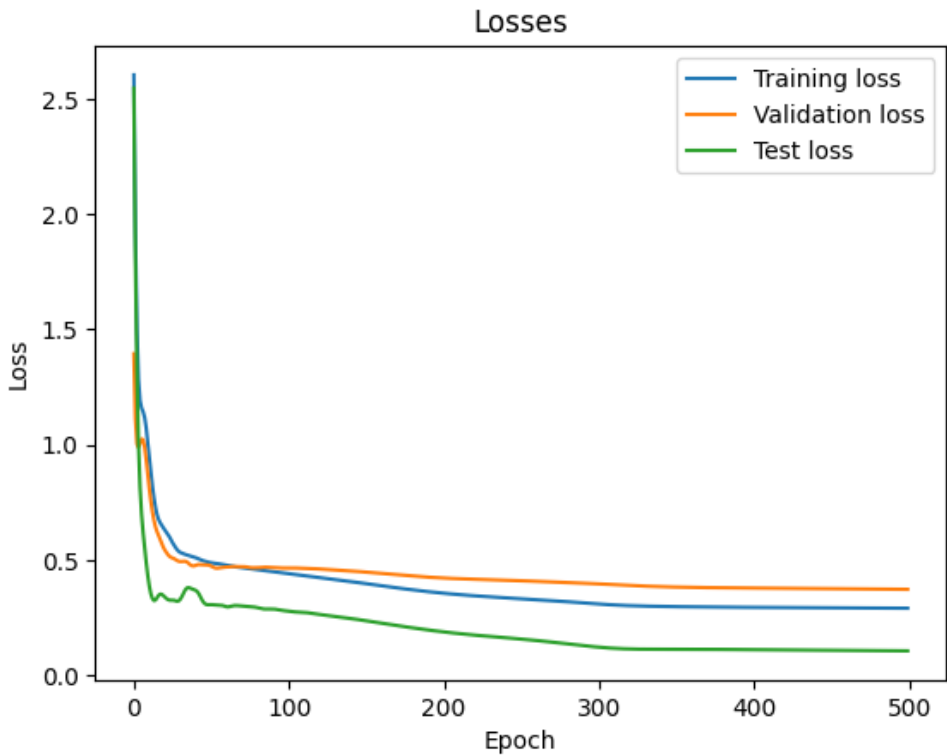
Метод обучения	Скорость обучения α	Режим обучения	Функция потерь
GD	0.04	batch	Quadratic loss

Метод инициализации сети: инициализация Хавьера

Критерий обучения:

$$E(w) = \frac{1}{n} \sum_{i=1}^n E^{(i)}(w) \rightarrow \min_w$$





Показатели качества обученной нейросетевой модели:

	Обучающая	Валидационная	Тестовая
Макс. абс. ошибка	0.390	0.463	0.222
С.к.о. ошибок	0.292	0.373	0.107
RMSE	0.540	0.611	0.327

Выводы: При стохастическом режиме обучения, изменение весов слишком сильное и чувствительное к конкретным точкам. Из-за этого loss функция не имеет свойств гладкости. В связи с этим, был выбран batch режим обучения. Можно заметить, что после 300 эпох изменений в loss функциях нет, то есть можно остановить обучение. Также, при слишком большом количестве нейронов (40), они дублируют друг друга, в связи с тем было уменьшено их количества до 10.