

Московский государственный университет им. М. В. Ломоносова

Физический факультет

Кафедра математики

**Оптимизация архитектуры линейного персептрона с помощью
генетического алгоритма.**

Курсовая работа по физике
студента 206 группы
Клиентова Григория Алексеевича

Научный руководитель:
д. ф.-м. н., профессор П. В. Голубцов

Москва

2022

Оглавление

1. Введение.....	3
2. Цель работы.....	3
3. Рассмотренные виды задач для линейного персептрона.....	3
4. Описание работы генетического алгоритма.....	4
4.1 Описание среды.....	4
4.2 Описание особей.....	4
4.3 Описание способа мутации.....	5
5. Описание реализованного алгоритма.....	5
6. Полученные результаты.....	6

1. Введение

Искусственные нейронные сети (ИНС) всё чаще начинают использоваться для решения различных прикладных задач широкой разновидности. Они позволяют находить субоптимальные решения за адекватное время и с использованием разумных ресурсов в условиях неполной или недостаточной априорной информации для построения точного решения.

Однако выбор в архитектуре модели, использующейся для нахождения такого решения может приводить к довольно различающимся результатам. В подавляющем числе случаев вид ИНС подбирают эмпирическим методом — путём перебора некоторых вариантов моделей нейронных сетей.

В данной работе рассмотрен один из методов для выбора наилучшей архитектуры линейной нейронной сети — персептрона — с помощью применения генетического алгоритма.

2. Цель работы

- Рассмотреть метод оптимизации архитектуры персептрона с помощью генетического алгоритма.
- Сделать вывод о преимуществе данного метода для выбора оптимальной архитектуры персептрона.

3. Рассмотренные виды задач для линейного персептрона

Для проверки работоспособности нейронных сетей и оценивания в первом приближении были использованы следующие «де факто» эталонные задачи:

- «Исключающее ИЛИ».
- Аппроксимация некоторых функций одного переменного

4. Описание работы генетического алгоритма

В рамках генетического алгоритма существуют особи — каждой из которой сопоставляется генотип, в котором зашифрована информация об архитектуре НС. На основе генотипа строится модель НС. Далее эта модель обучается на тренировочных данных, идентичных для каждой особи. После того, как особи были обучены на определённом количестве эпох происходит валидация каждой особи на специальном валидационном наборе данных. Каждой особи присваивается значение ошибки, которое получилось на этом наборе. Далее происходит эволюционный шаг — среда, в которой содержится популяция, генерируют новый набор особей на основе предыдущей популяции путём мутации и с учетом стратегии элитизма — особи, показавшие наилучший результат, переходят в новую популяцию без мутаций генов.

4.1 Описание среды

Была выбрана модель среды, которая содержит в тренировочный и валидационный наборы, а также популяцию особей, число которых на каждом эволюционном шаге константно. Изначально генотип каждой особи генерируется случайным образом.

4.2 Описание особей

Особь содержит в себе полносвязную нейронную сеть. Информация о её структуре содержится в генотипе в виде последовательного списка слоев. Слой может быть одним из списка:

- Линейный слой, содержащий нейроны
- Слой активации с функцией ReLU
- Слой активации с функцией Tanh
- Слой активации с функцией Sigmoid

4.3 Описание способа мутации

Мутация происходит на основе генотипа одной особи. 3 особи,

показавшие наилучшую, то есть наименьшую, ошибку на валидационном наборе дублируются в новую популяцию без изменений, однако у них заново пересоздается модель нейронной сети, чтобы обновить веса для обучения на новой эволюционной эпохе. В новую популяцию также входят особи, мутировавшие из этих трёх. Такая стратегия называется стратегией элитизма. Она призвана улучшать сходимость генетического алгоритма.

Три особи, показавшие самые наихудшие ошибки на валидационном наборе, не проходят в новую популяцию, даже без мутаций.

У особей, которые занимают промежуточные положения в ранжированном списке ошибок на валидационном наборе, переходят в новую популяцию только потомки, прошедшие мутации. Каждая мутация может происходить с заранее определенной вероятностью.

Возможны следующие виды мутации:

1. Изменение количества входных параметров у скрытого слоя нейронов.
2. Добавления промежуточного скрытого слоя
3. Замена слоя активации на слой нейронов
4. Удаление слоя

5. Описание реализованного алгоритма

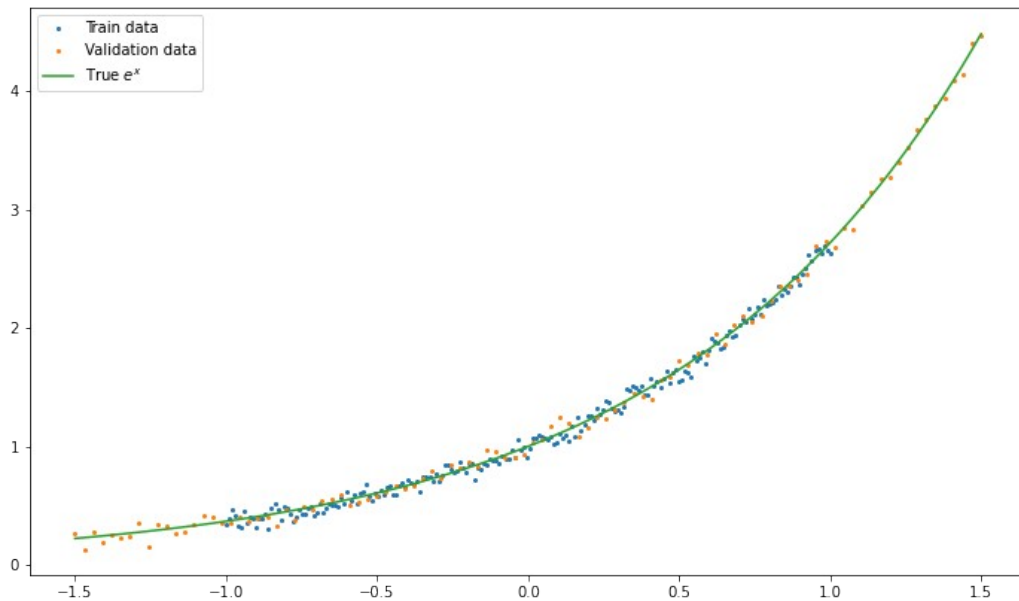
Изначально создается популяция, генотипы особей которых генерируются случайным образом. Далее эта популяция обучается на тренировочном наборе. Все наборы данных одинаковы для любых особей любых популяций. После обучения на заданном количестве обучающих эпох происходит валидация каждой особи в популяции и определение их ошибок на этом наборе. На основе этих данных происходит эволюционный шаг — в данном случае мутация особей по ранее описанному алгоритму.

6. Полученные результаты

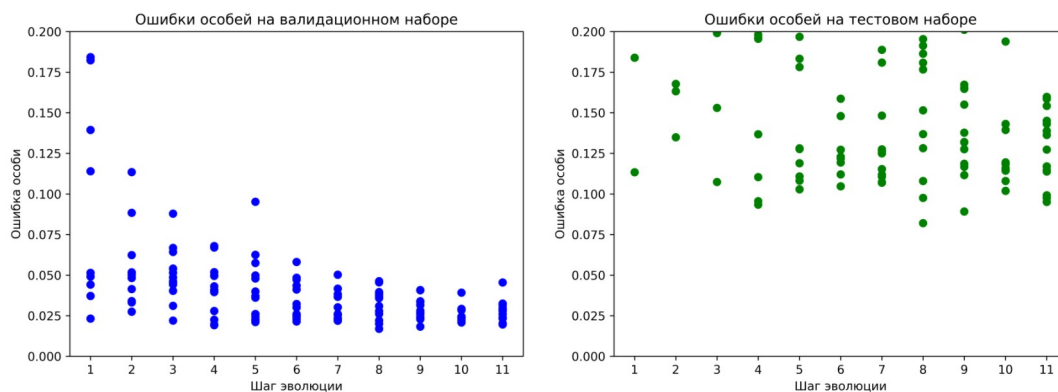
6.1 Задача аппроксимации

Поставим задачу аппроксимации функции.

Возьмем в качестве исходной функции график $y=e^x$ с шумом.



Создадим 15 особей со случайным набором генов. Установим количество обучающих эпох равным 25 на каждую эволюционную эпоху. Спустя 11 эволюционных эпох посмотрим на результаты обучения особей.

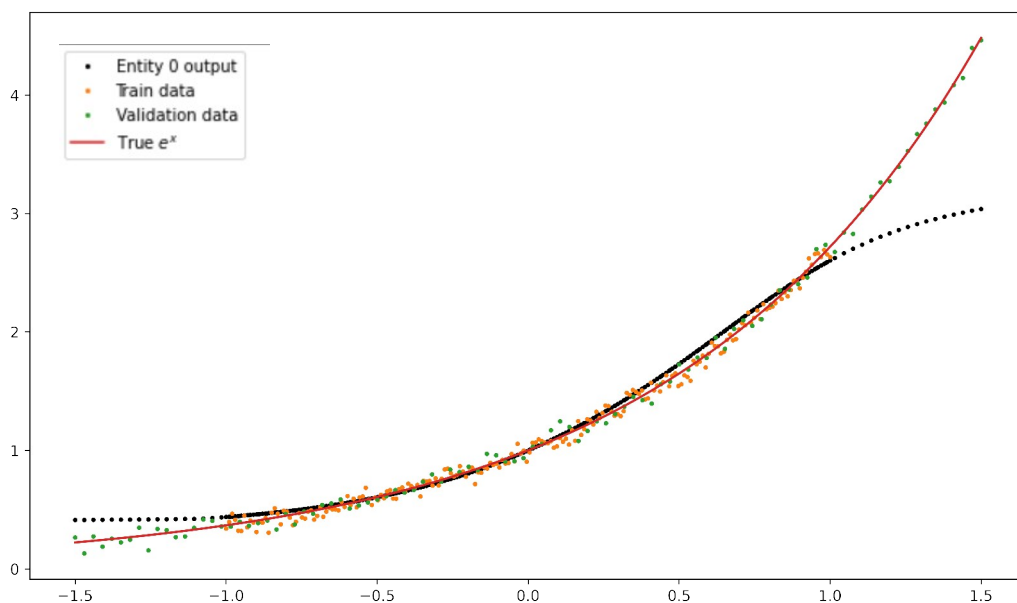


Как можно видеть из графика ошибки на валидационном наборе, размах ошибок среди всех особей на одной и той же эволюционной эпохе уменьшается, однако минимальная ошибка среди всех особей не сильно

отличается от минимума ошибки среди всех эпох и особей. Согласно стратегии элитизма минимальная ошибка среди особей должна не уменьшаться с течением эволюционных эпох, так как особи, показавшие наилучший результат, переходят в новую популяцию без изменения архитектуры. Однако из-за того, что при переходе в новую популяцию все особи заново инициализируют, причем случайным образом, обучение идентичных особей различается, как и получавшаяся ошибка.

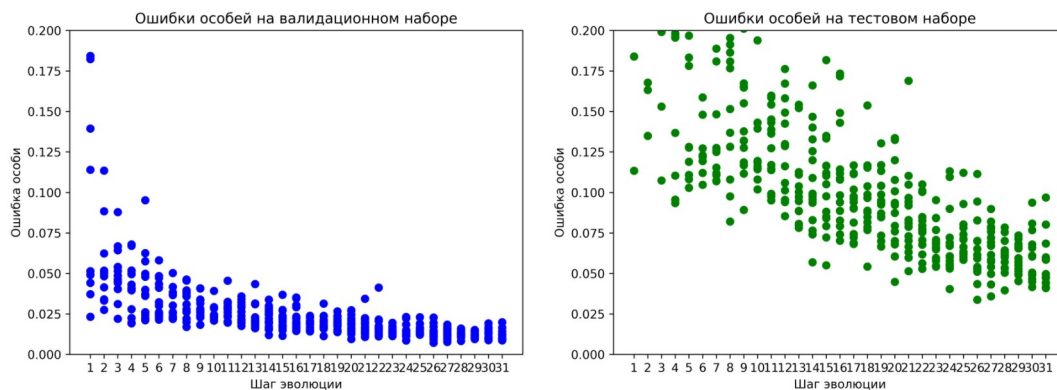
Из графика ошибки на тестовом наборе видно, что её размах внутри одной популяции также уменьшается и с количеством эволюционных эпох всё больше особей ближе к минимальной ошибке.

Посмотрим на результат одной из особей в последней популяции.



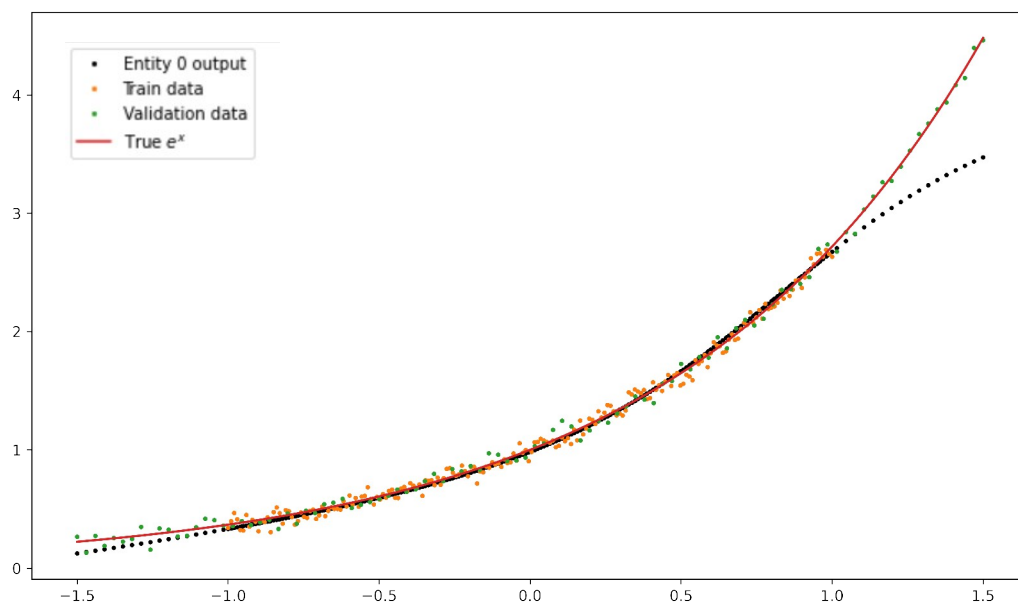
Как видно, сеть особи уже хорошо аппроксимирует исходную функцию, однако интерполяция получается неудовлетворительная.

Попробуем обучить особей на ещё 20 эволюционных эпохах.



Как видно из графиков, ошибки на тестовом и валидационном наборах падают, что говорит о работоспособности генетического алгоритма на данной задаче.

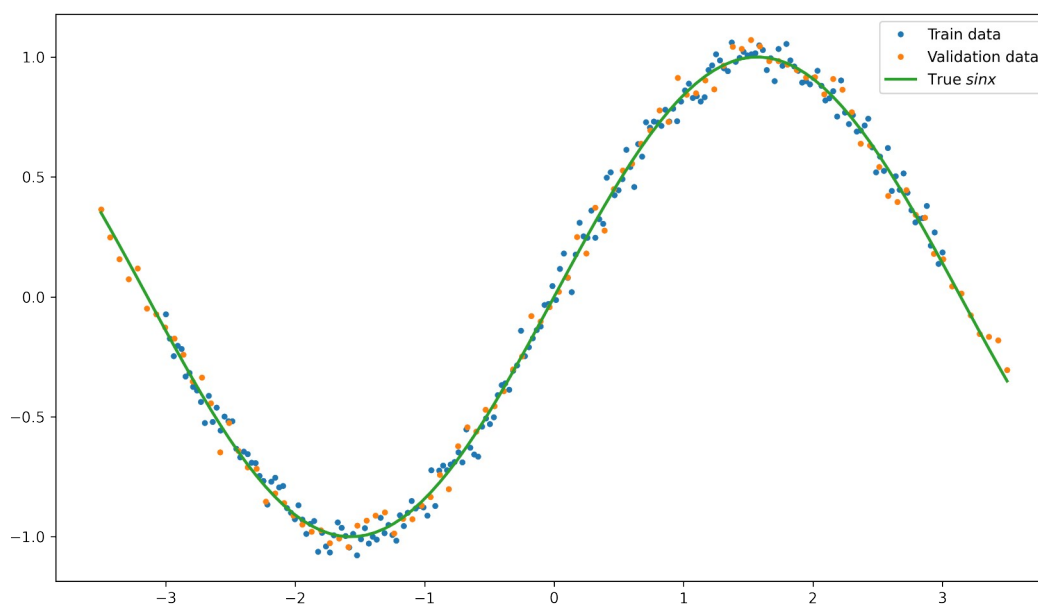
Посмотрим на предсказание одной из особей на последней популяции.



Можно увидеть, что особь научилась очень хорошо аппроксимировать исходную функцию, даже с учетом шума, однако интерполяция все также неудовлетворительная.

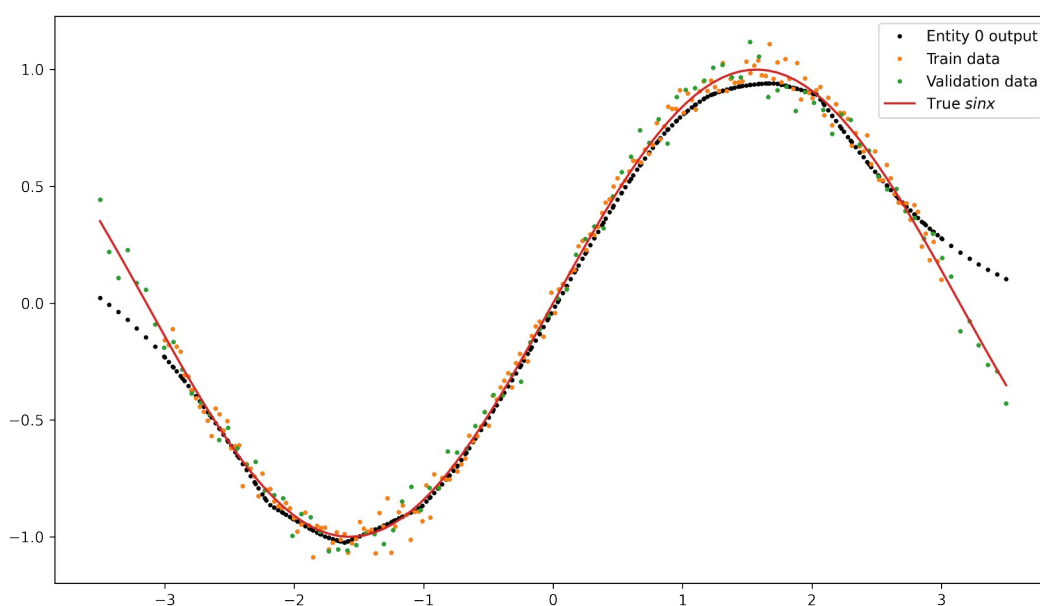
Попробуем выбрать в качестве исходной функции график вида $y = \sin x$.

Настройки генетического алгоритма идентичны.

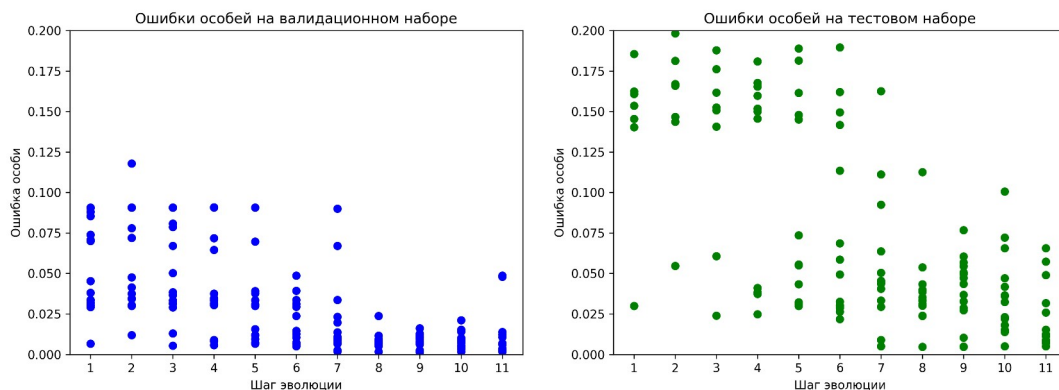


После обучения на 11 эволюционных эпохах были получены следующие результаты.

Предсказание одной из особей из последней популяции:
Как можно увидеть, даже после 10 эволюционных эпох

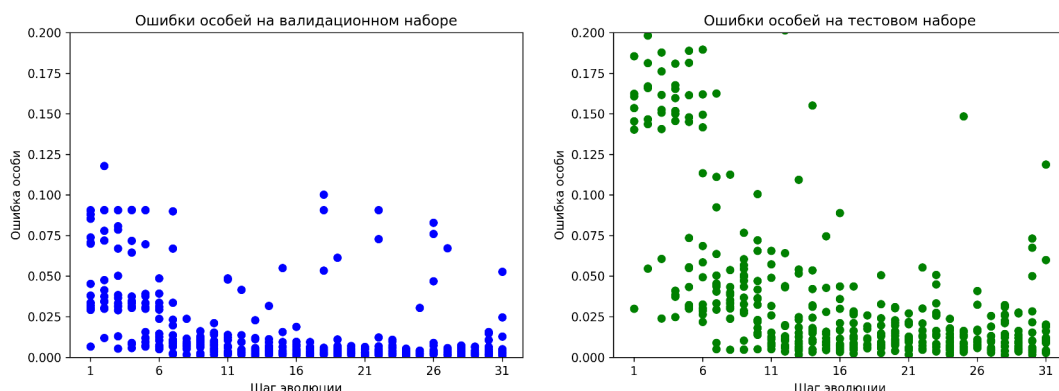


аппроксимация исходной функции достаточно хорошая.



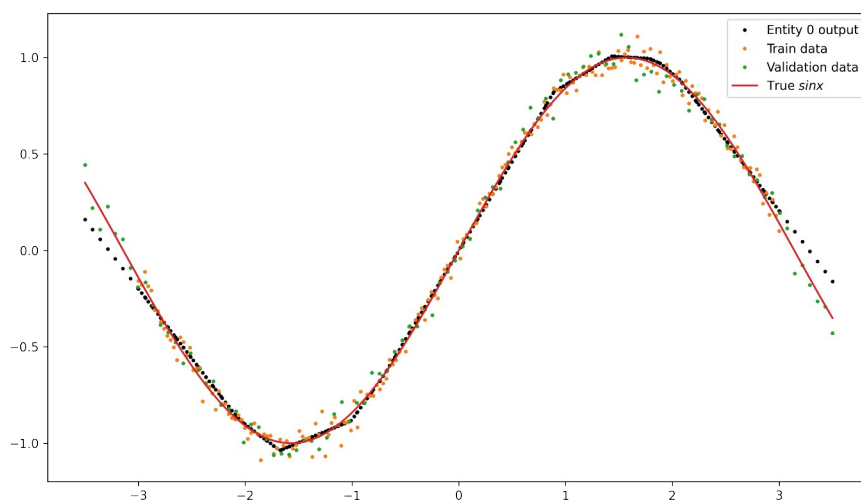
При близком рассмотрении графиков ошибок можно заметить, что размах ошибок особей внутри одной популяции уменьшается с шагом эволюции, как и при выборе экспоненты в качестве исходной функции.

Посмотрим на результаты спустя ещё 20 эволюционных эпох.



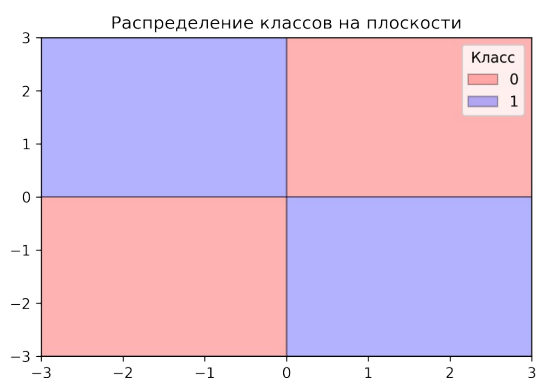
Как можно увидеть, ошибки сети уже после 10 эволюционных эпох сильно близки к 0.

Предсказание одной из особей из последней популяции:



6.2 Задача классификации «Исключающее ИЛИ»

Поставим задачу классификации точек на плоскости. Распределение классов показано следующим образом:



Исходя из этих областей генерируем данные для тренировочного, валидационного и тестового наборов.

Сами наборы выглядят следующим образом: