

Содержание

1	Введение	2
2	Постановка задачи	3
2.1	Описание данных и постановка задачи	3
2.2	Способ представления агностической неронной сети	3
2.2.1	WANN №1	4
2.2.2	WANN №2	4
3	Метрики	4
3.1	Эталонные метрики	4
3.2	Структурные метрики	5
3.3	Методы сравнения метрик	6
4	Генетический алгоритм и его модификации	7
4.1	Общие понятия	7
4.1.1	WANN 1	8
4.1.2	WANN 2	8
4.2	Классический генетический алгоритм	8
4.3	Детектирование стагнаций с помощью метрики	9
4.4	Детектирование стагнаций по значению ошибок	9
4.5	Генетический алгоритм с кластеризацией	10
5	Вычислительный эксперимент	10
5.1	Метрики	10
5.2	Генетические алгоритмы	13
6	Заключение	17
7	Список литературы	17

Аннотация

В работе исследуется задача аппроксимации неизвестной зависимости с помощью агностических нейронных сетей. На множестве агностических нейронных сетей вводится структурная метрика. Предлагается ряд эталонных метрик, а так же методы сравнения метрик друг с другом. На искусственно сгенерированных данных производится сравнение метрик. Рассматриваются четыре вариации генетического алгоритма, используемые для генерации агностических нейронных сетей.

Ключевые слова: WANN, структурное обучение, метрика, генетический алгоритм.

1 Введение

Актуальность темы. Во многих прикладных задачах требуется восстановить неизвестную зависимость, содержащуюся в данных. Одну из основных сложностей составляет именно отыскание наилучшей модели.

Цель работы. Ввести структурную метрику на множестве агностических нейронных сетей, предложить модификацию генетического алгоритма для генерации агностических нейронных сетей.

Методы исследования. Весь необходимый для вычислительного эксперимента код был написан на языке Python, в качестве примера были взяты данные из датасета "Iris dataset".

Научная новизна.

- Предложены методы сравнения метрик, заданных на конечных множествах.
- Рассмотрены структурные метрики на множестве агностических нейронных сетей.
- Изучены четыре модификации генетического алгоритма для генерации агностических нейронных сетей.

Практическая ценность. Написан класс агностических нейронных сетей, реализованы методы для сравнения метрик на конечном множестве, реализованы модификации генетических алгоритмов.

Обзор литературы. Одним из методов для решения задачи восстановления функциональной зависимости по набору исходных данных является символьная регрессия [1]. Задача символьной регрессии заключается в отыскании такой композиции функций из некоторого заданного множества, которая бы наилучшим образом аппроксимировала зависимость. Джон Коза предложил метод, являющийся аналогом эволюционного алгоритма [2], получивший дальнейшее развитие в работе Ивана Зелинки [3]. Анна Варфоломеева для решения задачи символьной регрессии использует совершенно другой подход, строящий дерево композиции на основе уже имеющихся примеров [4].

Основной интерес для данной работы представляет статья Андрея Кулунчакова, в которой он предлагает метод борьбы со стагнациями генетического алгоритма для генерации регрессионных деревьев. На множестве регрессионных деревьев вводится структурная метрика. В качестве эталонной метрики, которой должна соответствовать структурная, берется следующая:

$$\eta(f, f') = \frac{1}{|C|(|C| - 1)} \sum_{d_j, d_k \in C} [f(d_j) < f(d_k)][f'(d_j) > f'(d_k)]$$

Здесь C - некоторое наперед заданное множество. Кулунчаков предлагает считать метрики μ , η близкими на множестве деревьев \mathfrak{M} , если

$$\alpha(\mathfrak{M}) = \nu([\mu(f, f') \leq \alpha_1] \rightarrow [\eta(f, f') \leq \alpha_2] | f, f' \in \mathfrak{M}) \geq 1 - \varepsilon$$

где $\alpha_1, \alpha_2, \varepsilon$ - константы, $\nu(A)$ - частота события A . В нашей работе не используется такой способ сравнения метрик, потому что он требует задания сразу трех констант, что чрезмерно.

Далее Кулунчаков предлагает детектировать стагнации с помощью структурной метрики, считая, что алгоритм застagnировался в случае, если диаметр популяции ниже некоторого фиксированного порога, то есть, иными словами, если все деревья в популяции слишком похожи друг на друга. В этом случае половина худших деревьев в популяции меняется на заново сгенерированные.

В данной работе схожий подход был применен для агностических нейронных сетей (Weight Agnostic Neural Networks), подробно описанных в статье [5]. Суть данной модели заключается в том, что все веса на ребрах нейронной сети заменяются одним единственным весом, после чего подбирается такая структура, при которой агностическая нейронная сеть показывает наилучшие результаты в некотором диапазоне значений этого единственного параметра.

2 Постановка задачи

2.1 Описание данных и постановка задачи

Задана пара (X, y) , $X \in \mathbb{R}^{n \times m}$, $y \in \mathbb{R}^{n \times k}$, строки матриц X и y находятся в некоторой неизвестной функциональной зависимости. Задано множество порождающих функций \mathfrak{F} . $\mathfrak{F}_{\mathfrak{G}}$ - множество агностических нейронных сетей, у которых функции активации содержатся в \mathfrak{G} .

Требуется найти оптимальную модель $f : \mathbb{W} \times \mathbb{X} \rightarrow \mathbb{R}^{n \times k}$, $f \in \mathfrak{F}_{\mathfrak{G}}$, \mathbb{W} - множество параметров модели (в случае агностической нейронной сети состоит из одного элемента).

Более формально, если задана функция ошибки $S(y_{pred}, y_{true})$, то необходимо найти $f \in \mathfrak{F}_{\mathfrak{G}}$ и $w \in \mathbb{W}$, минимизирующие значение функции S :

$$(f, w) = \arg \min_{(f, w) \in \mathfrak{F}_{\mathfrak{G}} \times \mathbb{W}} S(f^w(x), y)$$

Здесь и далее запись $f^w(x)$ эквивалентна $f(w, x)$.

В качестве функции ошибки используется среднее квадратов регрессионных остатков:

$$S(y_{pred}, y_{true}) = \frac{1}{n} \|y_{pred} - y_{true}\|_2^2$$

2.2 Способ представления агностической нейронной сети

В данной работе рассматриваются два способа представления агностических нейронных сетей. В дальнейшем нейросети, представленные первым способом, мы будем для краткости называть **WANN №1**, нейросети, представленные вторым способом - **WANN №2**. Если номер не указывается,

это означает, что сказанное не зависит от способа представления агностической нейронной сети.

2.2.1 WANN №1

Агностическая нейронная сеть в данном подходе представляется n слоями, причем могут быть соединены между собой только нейроны соседних слоев. В каждом слое обязательно присутствует bias. Иными словами, рассматривается неполносвязная сеть прямого распространения, у которой все параметры равны между собой.

2.2.2 WANN №2

У агностической нейронной сети в данном подходе присутствуют только входной и выходной слои. Остальные нейроны могут соединяться между собой в произвольном порядке, единственное требование - в графе конфигураций не должно быть направленных циклов.

3 Метрики

3.1 Эталонные метрики

Напомним основополагающее определение.

Definition 3.1. Функция $\mu : X \times X \longrightarrow \mathbb{R}_{\geq 0}$ называется метрикой, если она удовлетворяет следующим условиям:

1. $\mu(x_1, x_2) = 0 \iff x_1 = x_2$
2. $\mu(x_1, x_2) = \mu(x_2, x_1)$
3. $\mu(x_1, x_2) + \mu(x_2, x_3) \geq \mu(x_1, x_3)$ (неравенство треугольника)

В общем случае эталонной метрикой на множестве X будем называть метрику, которая постулируется исходя из каких-то эмпирических соображений. Такие метрики нужны для того, чтобы сравнивая с ними другие метрики понимать, насколько последние близки к нашим ожиданиям.

Важно отметить, что для эталонной метрики допустимо невыполнение первого свойства ($\mu(x, y) = 0 \iff x = y$). В таком случае мы просто считаем, что метрика отождествляет x и y , для нее они неразличимы. Можно еще сказать, что метрика разбивает наше множество на классы эквивалентности и мы отныне работаем только с их представителями.

Ниже приведен список метрик, которые в данной статье считаются эталонными для множества агностических нейронных сетей:

$$1) \mu_1(f_1, f_2) = \frac{1}{n|W|} \sum_{w \in W} \sum_{i=1}^n \|f_1^w(x_i) - f_2^w(x_i)\|_2$$

$$2) \mu_2(f_1, f_2) = \max_{w_1, w_2 \in W} \frac{1}{n} \sum_{i=1}^n \|f_1^{w_1}(x_i) - f_2^{w_2}(x_i)\|_2$$

Здесь $x_i, x_j \in X$, $X \subset \mathbb{R}^n$ - некоторый конечный набор точек, W - конечное множество параметров. Заметим, что эти метрики определены для агностических нейронных сетей, чьи входные и выходные размерности совпадают.

Lemma 3.1. *Все предложенные функции являются метриками.*

Доказательство. Выполнение всех свойств, кроме неравенства треугольника, очевидно. Докажем, что и последнее тоже выполнено.

1) Следует из того, что для каждого фиксированного w неравенство выполнено, так как при фиксированном w имеем не что иное как сумму евклидовых норм.

2) Пусть максимум для (f_1, f_3) достигается на (w'_1, w'_2) . Тогда $\sum_{i=1}^n \|(f_1^{w'_1}(x_i) - f_2^{w'_2}(x_i))\|_2 \leq$ (по свойствам евклидовой метрики) $\leq \sum_{i=1}^n \|(f_1^{w'_1}(x_i) - f_2^{w'_2}(x_i))\|_2 + \sum_{i=1}^n \|(f_2^{w'_2}(x_i) - f_3^{w'_2}(x_i))\|_2 \leq \max_{w_1, w_2 \in W} \sum_{i=1}^n \|(f_1^{w_1}(x_i) - f_2^{w_2}(x_i))\|_2 + \sum_{i=1}^n \|(f_2^{w_1}(x_i) - f_3^{w_2}(x_i))\|_2$ \square

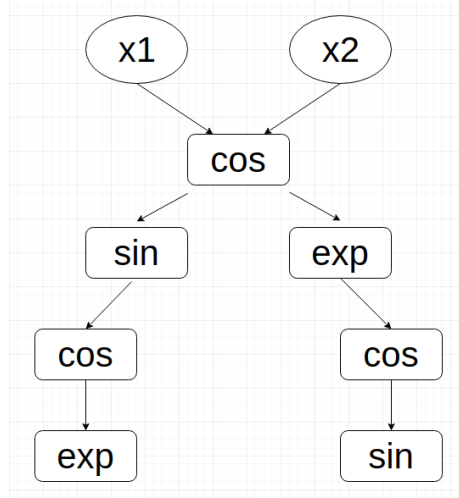
3.2 Структурные метрики

Структурной метрикой мы будем называть метрику, которая определена на множестве агностических нейронных сетей и зависит только от их структуры. Важно подчеркнуть, что, вообще говоря, вводится не метрика, а мера близости, метрикой мы ее будем называть лишь для лаконичности изложения. Из свойств метрики можно гарантировать лишь неотрицательность и симметричность. $\mu(x, y) = 0 \iff x$ изоморфен y - означало бы решить задачу проверки графов на изоморфизм, принадлежности которой классу P до сих пор не доказана. $\mu(x, y) = 0 \iff x$ и y ведут себя одинаково, как функции - означало бы решить задачу SAT, которая является NP-полной.

Напомним, что \mathfrak{F} - множество примитивных функций. Каждой агностической нейронной сети мы поставим в соответствие матрицу размера $(|\mathfrak{F}| + ar(f)) \times |\mathfrak{F}|$, где $ar(f)$ - арность функции, или, что то же самое, размер входного слоя сети. Строкам соответствуют все функции из \mathfrak{F} , а так же $x_1, \dots, x_{ar(f)}$, столбцам - только функции из \mathfrak{F} .

В ячейке с координатами (i, j) стоит число, равную количеству направленных ребер в агностической нейронной сети, ведущих из нейрона с функцией активации соответствующей i -ой строке в нейрон с функцией активации, соответствующей j -ому столбцу (определение аналогично для $x_1, \dots, x_{ar(f)}$).

Например, пусть нейросеть имеет следующий вид:



Тогда ей соответствует матрица

$$\begin{bmatrix} & \sin & \cos & \exp \\ \sin & 0 & 1 & 0 \\ \cos & 2 & 0 & 2 \\ \exp & 0 & 1 & 0 \\ x_1 & 0 & 1 & 0 \\ x_2 & 0 & 1 & 0 \end{bmatrix}$$

Первый способ измерять расстояние между двумя агностическими нейронными сетями по построенным матрицам - вытягивать матрицы в вектор и брать евклидову норму разности этих двух получившихся векторов.

Второй способ - подавать модуль разности векторов на вход нейронной сети, предварительно обученной предсказывать одну из эталонных метрик.

3.3 Методы сравнения метрик

В данном разделе вводятся три способа сравнения метрик. Неформально говоря, мы хотим, чтобы структурная метрика была в некотором роде близка к эталонной.

Более строго, пусть дано фиксированное конечное множество T (в нашем конкретном случае $T \subset \mathfrak{F}_{\mathfrak{G}}$). Необходимо ввести функцию $\gamma(\mu_1, \mu_2)$, по значению которой можно было бы понять, насколько согласуются между собой метрики на данном множестве.

1 способ. Посчитать обычную корреляцию.

$$\text{corr}(\mu_1, \mu_2) = \frac{\overline{\mu_1 \mu_2} - \overline{\mu_1} \cdot \overline{\mu_2}}{(\overline{\mu_1^2} - \overline{\mu_1}^2)(\overline{\mu_2^2} - \overline{\mu_2}^2)}$$

Здесь $\overline{\mu_i} = \frac{1}{n(n-1)} \sum_{t_1, t_2 \in T} \mu_i(t_1, t_2)$

2 способ. $\nu(\mu(f_1, f_2) < \mu(f_1, f_3) \wedge \eta(f_1, f_2) > \eta(f_1, f_3))$, ν - частота события

3 способ. Напомним определение эквивалентных метрик.

Definition 3.2. Метрики μ и ν называются эквивалентными, если

$$\exists \alpha, \beta \quad \forall x, y \quad \alpha \nu(x, y) \leq \mu(x, y) \leq \beta \nu(x, y).$$

В дальнейшем, говоря про α и β , мы имеем в виду наибольшее среди всех α , удовлетворяющих данному неравенству, и наименьшее среди всех таких β .

На конечных множествах все метрики эквивалентны. Хотим рассмотреть $g(\alpha, \beta)$, которая, с одной стороны, показывала бы, насколько α и β близки друг к другу (это и определяет меру близости наших метрик), с другой стороны, была бы устойчива к масштабированию метрик (т.е. не менялась бы при умножении метрик на константу). Одной из таких функций является $\frac{\alpha}{\beta}$, ее и будем рассматривать. Две метрики тем ближе, чем ближе значение этой функции к единице.

4 способ. Пусть для множества T заданы матрицы попарных расстояний D_1, D_2 для метрик μ_1 и μ_2 соответственно. Нормализуем эти матрицы (эквивалентно делению каждого элемента на диаметр множества T), вытянем каждую матрицу в вектор и будем смотреть на евклидово расстояние между этими векторами.

4 Генетический алгоритм и его модификации

4.1 Общие понятия

Напомним, что нашей исходной задачей являлось отыскание пары (f, w) , т.ч.

$$(f, w) = \arg \min_{(f, w) \in \mathcal{F}_{\mathcal{G}} \times \mathcal{W}} S(f^w(x), y)$$

Хотя размерность пространства параметров в агностических нейронных сетях и так снижена до 1, все равно перебор по сетке может занимать слишком много времени. Поэтому мы рассмотрим три подхода к данной проблеме.

1 подход. На каждой итерации генетического алгоритма рассматривается усредненное значение ошибки по разреженной сетке (в данной работе -2, -1, -0.5, 0.5, 1, 2). Таким образом ищется нейросеть, именно структура которой хорошо подходит под решение задачи, а параметр слабо влияет на результат.

2 подход. На каждой итерации генетического алгоритма рассматривается максимальное значение ошибки по разреженной сетке. Цель та же, что и у предыдущего подхода.

3 подход. На каждой итерации генетического алгоритма рассматривается минимальное значение ошибки по разреженной сетке. Такой подход ближе в первоначальной постановке задачи.

4.1.1 WANN 1

Мутация. Операция мутации производит 3 возможных действия:

- Удаляет случайно выбранный слой, соседние с ним слои соединяет между собой произвольным образом.
- Заменяет случайно выбранный слой на новый, соединяет с соседними слоями произвольным образом.
- В случайное место добавляет новый сгенерированный слой, соединяет с соседними слоями произвольным образом.

Скрещивание. Операция скрещивания:

1. Случайным образом выбирает в первой сети и второй сети по одному слою.
2. Все слои, которые шли в первой сети после выбранного в ней слоя, заменяет на слои второй сети, начиная с выбранного.

В следующем примере продемонстрирована работа алгоритма скрещивания:

$$L_1^1 L_2^1 \dots L_k^1 \dots L_s^1, L_1^2 L_2^2 \dots L_m^2 \dots L_r^2 \longrightarrow L_1^1 \dots L_k^1 L_m^2 \dots L_r^2$$

Здесь L_i^j - i -й слой j -й нейросети.

4.1.2 WANN 2

Мутация. Операция мутации производит 3 возможных действия:

- Добавляет новое ребро.
- Добавляет новый нейрон на уже существующее ребро.
- Изменяет функцию активации у уже существующего нейрона.

Скрещивание. Отсутствует.

4.2 Классический генетический алгоритм

В классической вариации генетического алгоритма ведется ненаправленный поиск оптимальной структуры нейросети с помощью применения к популяции описанных ранее операциях скрещивания и мутации.

Result: WANN

```
generate_population();  
while error >  $\varepsilon$  or iteration < max_iterations do  
    mutate();  
    crossover(); //for WANN №1;  
    select();  
end  
return population[0];
```

Algorithm 1: Классический генетический алгоритм

4.3 Детектирование стагнаций с помощью метрики

В статье Кулунчакова [4] предложено детектировать стагнации в работе генетического алгоритма с помощью метрики: если диаметр популяции (как множества) не превосходит некоторого ε , то мы считаем, что алгоритм застагнировался и меняем половину худших деревьев в популяции на совершенно случайные.

```
Result: WANN
generate_population();
while  $error > \varepsilon$  or  $iteration < max\_iterations$  do
    mutate();
    crossover(); //for WANN №1;
    select();
    if  $diam(population) \leq \delta$  then
        for  $wann$  in  $population[\frac{n}{2} :]$  do
            regenerate( $wann$ );
        end
    end
end
return population[0];
```

Algorithm 2: Генетический алгоритм с детектированием стагнаций с помощью метрики

4.4 Детектирование стагнаций по значению ошибок

Если лучшее значение ошибки в популяции перестает существенно меняться, то мы считаем это стагнацией и действуем аналогично предыдущему пункту, т.е. заменяем половину худших деревьев в популяции на совершенно случайные.

```
Result: WANN
generate_population();
while  $error > \varepsilon$  or  $iteration < max\_iterations$  do
    mutate();
    crossover(); //for WANN №1;
    select();
    if  $\frac{error[iteration-l]}{error[iteration]} \geq \delta$  then
        for  $wann$  in  $population[\frac{n}{2} :]$  do
            regenerate( $wann$ );
        end
    end
end
return population[0];
```

Algorithm 3: Генетический алгоритм с детектированием стагнаций по значению ошибки

4.5 Генетический алгоритм с кластеризацией

В этой вариации генетического алгоритма, если ошибка перестает изменяться, то множество нейросетей в популяции кластеризуется по структурной метрике, а потом в каждом кластере все нейросети, кроме той, у которой ошибка среди них минимальна, заменяются на заново сгенерированные. Тем самым мы избавляемся от похожих нейросетей в популяции.

```

Result: WANN
generate_population();
while  $error > \varepsilon$  or  $iteration < max\_iterations$  do
    mutate();
    crossover(); //for WANN №1;
    select();
    if  $\frac{error[iteration-l]}{error[iteration]} \geq \delta$  then
        clusters := cluster(population);
        for  $cluster$  in clusters do
            for  $wann$  in  $cluster \setminus \{best\_in\_cluster\}$  do
                | wann.regenerate();
            end
        end
    end
end
return population[0];

```

Algorithm 4: Генетический алгоритм с кластеризацией

5 Вычислительный эксперимент

5.1 Метрики

Были рассмотрены 4 метрики: $\|\cdot\|_2$; $\|\cdot\|_1$; $\|\cdot\|_{-\infty}$; $\|\cdot\|_{\infty}$. Они были попарно сравнены на множестве $X \subset \mathbb{R}^3$ ($|X| = 100$), взятого из равномерного распределения на $[-10, 10]^3$, с помощью каждого из четырех методов, описанных ранее.

	$\ \cdot\ _2$	$\ \cdot\ _1$	$\ \cdot\ _{-\infty}$	$\ \cdot\ _{\infty}$
$\ \cdot\ _2$	1.00	0.97	0.54	0.94
$\ \cdot\ _1$	0.97	1.00	0.70	0.84
$\ \cdot\ _{-\infty}$	0.54	0.70	1.00	0.34
$\ \cdot\ _{\infty}$	0.94	0.84	0.34	1.00

Таблица 1: 1 способ

	$\ \cdot\ _2$	$\ \cdot\ _1$	$\ \cdot\ _{-\infty}$	$\ \cdot\ _{\infty}$
$\ \cdot\ _2$	0.00	0.08	0.33	0.12
$\ \cdot\ _1$	0.08	0.00	0.26	0.19
$\ \cdot\ _{-\infty}$	0.33	0.26	0.00	0.40
$\ \cdot\ _{\infty}$	0.12	0.19	0.40	0.00

Таблица 2: 2 способ

	$\ \cdot\ _2$	$\ \cdot\ _1$	$\ \cdot\ _{-\infty}$	$\ \cdot\ _{\infty}$
$\ \cdot\ _2$	1.00	0.58	0.00	0.58
$\ \cdot\ _1$	0.58	1.00	0.00	0.34
$\ \cdot\ _{-\infty}$	0.00	0.00	1.00	0.00
$\ \cdot\ _{\infty}$	0.58	0.34	0.00	1.00

Таблица 3: 3 способ

	$\ \cdot\ _2$	$\ \cdot\ _1$	$\ \cdot\ _{-\infty}$	$\ \cdot\ _{\infty}$
$\ \cdot\ _2$	0.00	7.77	34.15	9.58
$\ \cdot\ _1$	7.77	0.00	26.90	16.95
$\ \cdot\ _{-\infty}$	34.15	26.90	0.00	42.20
$\ \cdot\ _{\infty}$	9.58	16.95	42.20	0.00

Таблица 4: 4 способ

Как видим, все способы дают на этих метриках одинаковый результат: если μ_1 ближе к μ_2 , чем к μ_3 , относительно одного из способов, то то же самое мы наблюдаем и во всех остальных.

Затем нейросеть с тремя скрытыми слоями была обучена предсказывать эталонные метрики по построенной матрице, как это уже описывалось ранее.

Ниже нейросеть, обученная предсказывать эталонную метрику №1, обозначается **net1**, а нейросеть, обученная предсказывать эталонную метрику №2 - **net2**. В качестве обучающей выборки были взяты попарные расстояния между сгенерированными WANN'ами, в качестве тестовой - попарные расстояния между другими ста.

Далее представлены результаты для агностических сетей первого типа.

	standart 1	standart 2
structural 1	0.31	0.37
net1	0.12	0.22
net2	0.13	0.35

Таблица 5: 1 способ

	standart 1	standart 2
structural 1	0.44	0.46
net1	0.49	0.49
net2	0.52	0.51

Таблица 6: 2 способ

	standart 1	standart 2
structural 1	0.00	0.00
net1	0.01	0.03
net2	0.00	0.20

Таблица 7: 3 способ

	standart 1	standart 2
structural 1	21.92	29.01
net1	24.92	51.92
net2	22.79	27.92

Таблица 8: 4 способ

Таким образом, получаем соответствие между эталонной метрикой и той, которая лучше всего с ней согласуется.

	1 способ	2 способ	3 способ	4 способ
standard 1	structural 1	structural 1	net 1	structural 1
standart 2	structural 1	structural 1	net 2	net 2

Таблица 9: Соответствие между эталонными и структурными метриками

Приведем аналогичные результаты для агностических нейронных сетей второго типа.

	standart 1	standart 2
structural 1	0.32	0.34
net1	0.28	0.20
net2	0.18	0.28

Таблица 10: 1 способ

	standart 1	standart 2
structural 1	0.44	0.47
net1	0.44	0.47
net2	0.47	0.47

Таблица 11: 2 способ

	standart 1	standart 2
structural 1	0.00	0.00
net1	0.00	0.00
net2	0.00	0.08

Таблица 12: 3 способ

	standart 1	standart 2
structural 1	36.94	21.82
net1	23.12	45.04
net2	27.39	31.30

Таблица 13: 4 способ

	1 способ	2 способ	3 способ	4 способ
standard 1	structural 1	structural 1 / net 1	—	net 1
standart 2	structural 1	—	net 2	structural 1

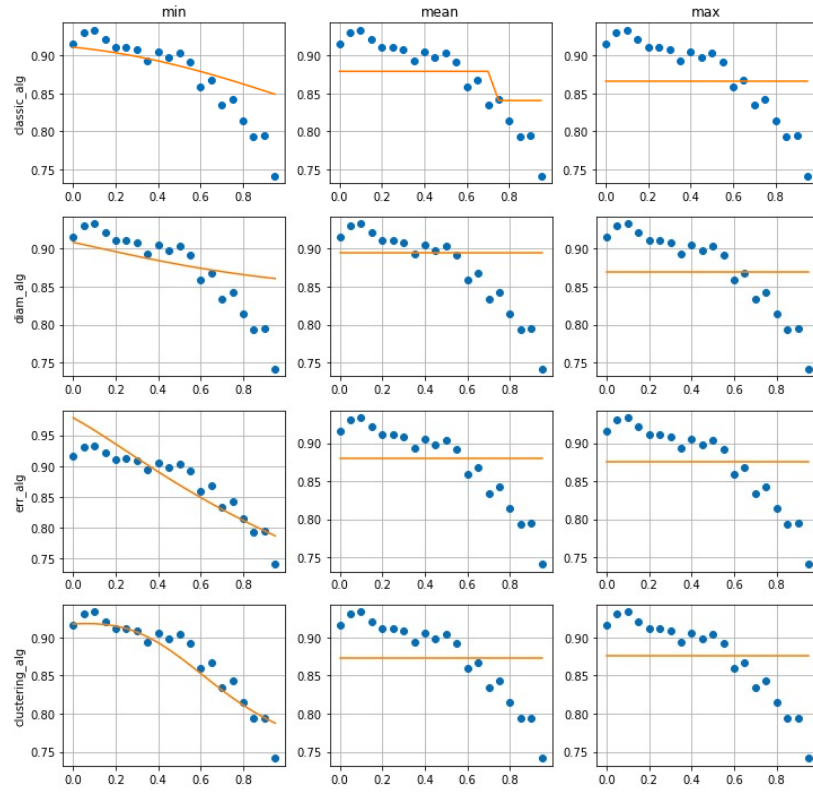
Таблица 14: Соответствие между эталонными и структурными метриками

Таким образом, предсказание эталонных метрик с помощью нейросетей на тестовой выборке показывает хоть и не слишком плохие результаты, но не лучшие, чем структурная метрика №1, поэтому в дальнейшем мы будем использовать именно последнюю.

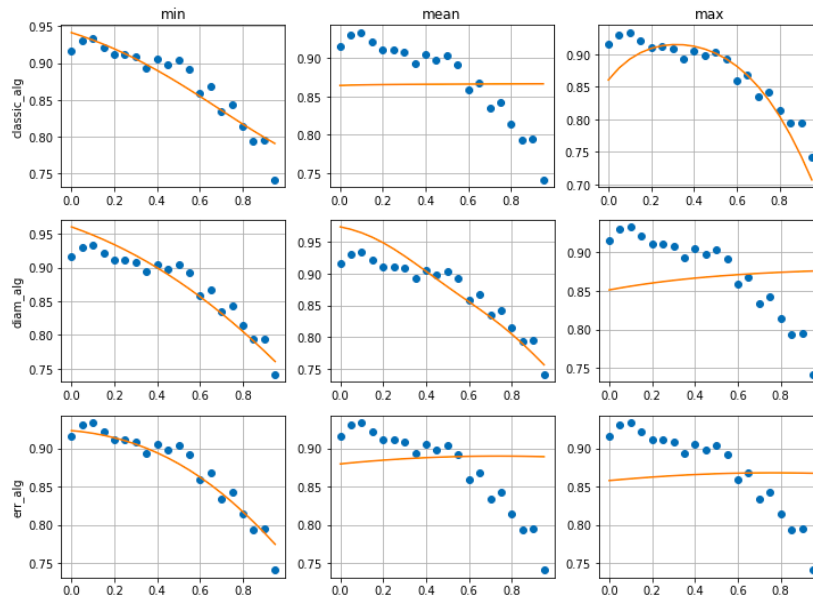
5.2 Генетические алгоритмы

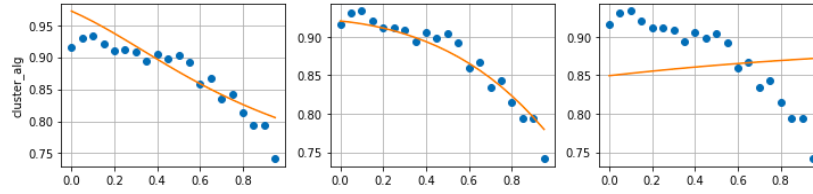
Сначала работа алгоритмов была протестирована на искусственно сгенерированных данных. Каждая из модификаций генетического алгоритма была запущена на 50 эпох с каждой из трех функций ошибки. Получившаяся агностическая нейронная сеть была оптимизирована по параметру. Ниже на графиках представлены исходное множество точек и их функциональная аппроксимация.

Для агностических сетей первого типа:

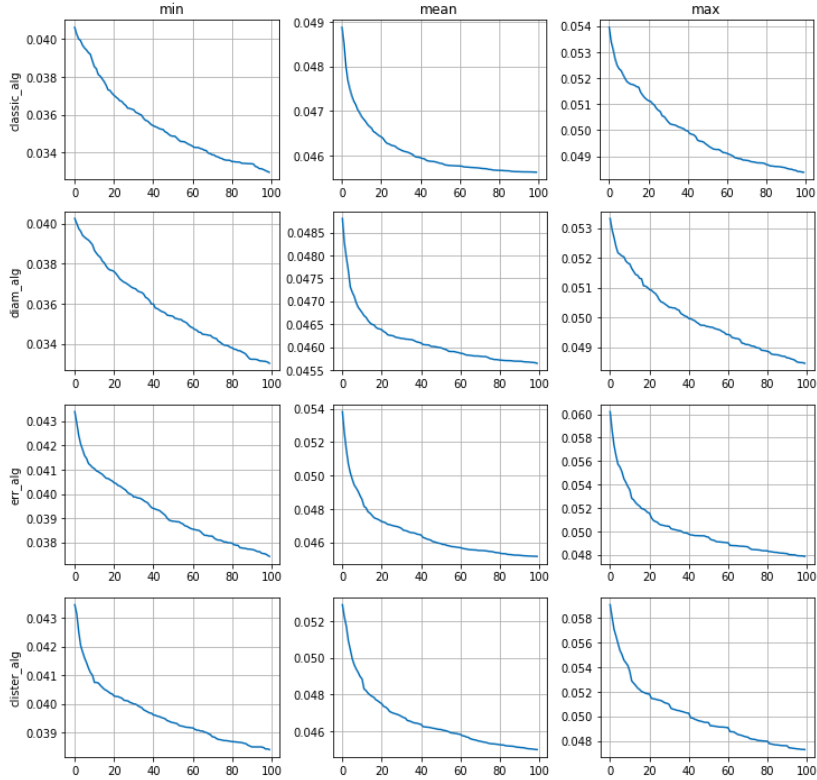


Для агностических сетей второго типа:





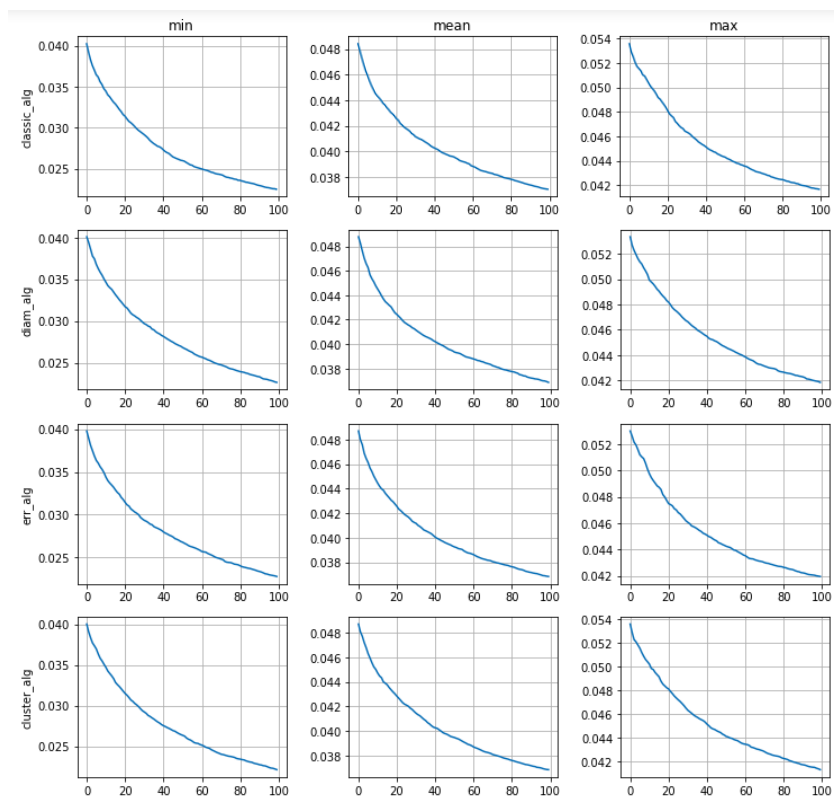
После этого то же самое было проделано для датасета Iris. Ниже на графиках приведена зависимость усредненного по экспериментам значения функции ошибки от числа эпох. Число проведенных экспериментов, по которым усреднялся результат - 100. Так же в таблицах представлено значение функции ассигасы на полученной агностической нейронной сети после оптимизации параметра. Оно так же было усреднено по всем экспериментам. Для агностических сетей первого типа:



	min	mean	max
classic	0.70	0.37	0.33
diameter	0.70	0.36	0.33
error	0.57	0.39	0.32
clustering	0.61	0.37	0.33

Таблица 15:

Для агностических сетей второго типа:



	min	mean	max
classic	0.88	0.77	0.64
diameter	0.87	0.77	0.63
error	0.88	0.81	0.62
clustering	0.88	0.80	0.65

Таблица 16:

6 Заключение

TODO

7 Список литературы

- [1] Riccardo Poli, William B. Langdon, Nicholas F. McPhee A Field Guide to Genetic Programming, 2008.
- [2] Koza, J. R. Genetic programming. In Williams, J. G. and Kent, A. (editors) // Encyclopedia of Computer Science and Technology, 1998. Vol. 39. P.: 29-43.
- [3] Ivan Zelinka, Zuzana Oplatkova, and Lars Nolle Analytic programming – symbolic regression by means of arbitrary evolutionary algorithms, August 2008.
- [4] А. А. Варфоломеева Выбор признаков при разметке библиографических списков методами структурного обучения, 2013.
- [5] <https://weightagnostic.github.io/>