

УДК 004.023

## Использование генетического алгоритма для выбора структуры искусственной нейронной сети

**Н.Е. Матюхин\*, А.А. Соколов**

Сибирский федеральный университет, пр. Свободный, 79,  
Красноярск, 660041, Россия

\*E-mail: nmatryukhin@mail.ru

**Аннотация.** Значительное внимание в статье уделяется исследованию возможности использования генетического алгоритма в задаче подбора структуры нейронной сети. Основное внимание уделяется проработке стадий эволюционного алгоритма и выбору наилучших комбинаций шагов алгоритма. Научная новизна заключается в использовании такого рода решения для поиска оптимальной структуры как задачи минимизации функции ошибки нейронной сети. В результате определена наилучшая конфигурация алгоритма, дающая значительный выигрыш при подборе аналогичной структуры методом полного перебора всех возможных решений.

**Ключевые слова:** генетический алгоритм, нейронные сети, структура сети, оптимизация

## Using a genetic algorithm to select the structure of an artificial neural network

**N.E. Matyukhin\*, A.A. Sokolov**

Siberian Federal University, 79 Svobodny pr., Krasnoyarsk, 660041, Russia

\*E-mail: nmatryukhin@mail.ru

**Abstract.** Considerable attention is paid to the study of the possibility of using the genetic algorithm in the problem of selecting the structure of a neural network. The main attention is paid to the elaboration of the stages of the evolutionary algorithm and the choice of the best combinations of the steps of the algorithm. Scientific novelty lies in the use of this kind of solution to find the optimal structure as a problem of minimizing the error function of a neural network. As a result, the best configuration of the algorithm was determined, which gives a significant gain when selecting a similar structure by the method of exhaustive enumeration of all possible solutions.

**Keywords:** genetic algorithm, neural networks, architecture of network, optimization

## 1. Введение

Искусственные нейронные сети – достаточно эффективный инструмент решения задач, сложных для решения алгоритмическим путем. Нейронные сети успешно применяются для решения задач распознавания образов, аппроксимации, прогнозирования, создания экспертных систем [3]. Они нашли широкое применение в нашей жизни, взяв на себя решение многих не самых простых задач [12]. Распознавание изображений, воспроизведение и перевод человеческой речи, обработка видеопотока в реальном времени – все эти задачи уже являются посильными для некоторых видов нейронных сетей. Они также находят своё применение в экономике и юриспруденции [8,11], а также позволяют максимально качественно изучать предпочтения миллионов пользователей и предлагать им ту рекламу, которая с высоким шансом может их заинтересовать [12]. Однако, для качественной работы таких сетей требуется решить ряд некоторых неспецифичных задач: подбор определенной структуры нейронной сети, соответствующей конкретной решаемой задаче, а также настройка весовых коэффициентов связей между нейронами сети.

Проблема подбора оптимальной архитектуры для конкретной нейронной сети составляет основную трудность при работе с ними [1]. Архитектура сети определяется числом нейронов на скрытых слоях, их активационными функциями, а также наличием связи между конкретными нейронами.

В данной работе для настройки структуры искусственной нейронной сети применяется генетический алгоритм. Также проведено сравнение данного подхода с полным перебором.

## 2. Подходы к подбору структуры нейронной сети

За все время развития нейросетевых технологий были разработаны различные подходы, позволяющие формализовать процесс подбора структуры нейронной сети: статистический, деструктивный и конструктивный.

Суть статистического подхода заключается в определении взаимосвязи заданных параметров ИНС на основе статистических методов [7]. В качестве параметров могут рассматриваться сложность ИНС, количество примеров в обучающей выборке, ошибка обучения, ошибка обобщения и другие [7].

Деструктивный подход заключается в изначальном выборе избыточной структуры нейронной сети, после чего происходит процесс упрощения [7]. Данный процесс заключается в постепенном удалении избыточных нейронов и связей, что

позволяет добиться оптимальной структуры. Недостатком данного подхода является произвольность выбора начальной структуры. В случае, если изначальная структура обладала недостаточной выразительностью для решения задачи, процесс ее упрощения не приведет к улучшению работы нейронной сети.

Суть конструктивного подхода заключается в начальном выборе минимальной структуры и ее дальнейшем наращивании путем добавления новых нейронов или дополнительных скрытых слоев. Теоретически это позволяет построить сеть с достаточно небольшим количеством слоев. Однако, на практике это вызывает свои сложности [2].

Наиболее эффективным подходом к решению задачи подбора структуры искусственной нейронной сети является оптимизационный.

### 3. Постановка задачи структурной оптимизации нейронной сети

Оптимизационный подход состоит в представлении задачи поиска оптимальной структуры как задачи минимизации функции ошибки нейронной сети. Оптимизируемыми параметрами в данном случае являются количество скрытых слоев, количество нейронов на скрытых слоях, а также функции активации:

$$E(S, W, B, TF) = \min_{S, TF}$$

Здесь  $S$  – матрица, определяющая наличие синаптических связей между элементами нейронной сети,  $W$  – матрица весовых коэффициентов, соответствующих связям,  $B$  – вектор смещений нейронов сети,  $TF$  – вектор функций активации нейронной сети.

Для решения поставленной задачи обычно применяются биоинспирированные метаэвристические алгоритмы поиска глобального экстремума. Несмотря на отсутствие доказательств о сходимости этих методов, их часто применяют для решения сложных оптимизационных задач. В данной статье для решения задачи настройки структуры нейронной сети используется генетический алгоритм. Применение ГА обусловлено тем, что генетический алгоритм обеспечивает глобальный просмотр пространства параметров, позволяя избегать локальных минимумов, а также не требует информации о градиентах.

#### **4. Эволюционное проектирование архитектуры нейронной сети**

В данной работе генетический алгоритм будет применен для решения задачи подбора архитектуры полносвязной искусственной нейронной сети прямого распространения.

Эволюционное проектирование архитектуры нейронной сети состоит из следующих шагов:

- декодирование каждой особи;
- обучение каждой особи методом обратного распространения ошибки;
- оценивание приспособленности каждой особи;
- репродукция особей согласно выбранному способу селекции;
- применение генетических операторов скрещивания и мутации для получения нового поколения.

#### **5. Обучение особи методом обратного распространения ошибки**

После создания хромосомы особи, необходимо оценить её приспособленность – показатель того, как успешно данная особь решает поставленную задачу. С целью нахождения этого параметра создаётся искусственная нейронная сеть с архитектурой, описанной в хромосоме особи. Полученную нейронную сеть необходимо обучить на тренировочной выборке данных. Для обучения применялся метод обратного распространения ошибки.

Метод обратного распространения ошибки был выбран за свою высокую эффективность в задаче обучения нейронной сети, а сам алгоритм не претерпевал изменений. Сначала входной сигнал распространяется через каждый слой, пока выходной слой не вернёт конкретное значение. После этого вычисляется ошибка распространения – разница между ожидаемыми выходами и выходами, полученными после прямого распространения. Эти ошибки затем распространяются по сети в обратном направлении и обновляют весовые коэффициенты.

#### **6. Оценивание приспособленности каждой особи**

Оценка функции приспособленности происходит путём вычитания из единицы значения средней квадратичной ошибки выхода полученной нейронной сети. Таким образом получается величина  $0 < \text{fitness} < 1$ , которая характеризует способность конкретной особи решать поставленную задачу. Чем ближе это значение к единице, тем

меньше ошибка, и как следствие, лучше себя показывает архитектура сети, предложенная данной особью.

## 7. Репродукция особей

Типы операторов селекции (турнирная, элитарная, рулеточная), рассматриваемые в данной работе обычно показывают себя одинаково при решении задачи подбора структуры искусственной нейронной сети [6]. В связи с этим было принято решение остановиться на гибриде рулеточного метода и элитарной селекции.

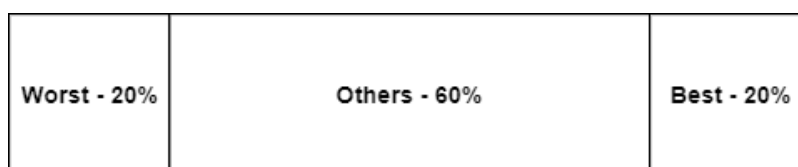
Как только было вычислено значение приспособленности каждой особи текущей популяции, определяется состав новой популяции.

Популяция предков (состоящая из особей-родителей) разделяется на три части по следующим правилам (рисунок 1):

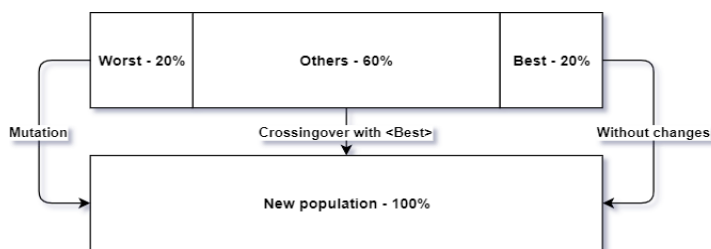
- 20% популяции с самым высоким значением приспособленности помечаются как «лучшие»;
- 20% популяции с самым низким значением приспособленности помечаются как «худшие»;
- оставшиеся 60% популяции помечаются как «другие».

После деления по значению приспособленности начинается формирование (селекция) популяции особей-потомков (рисунок 2):

- все «лучшие» особи переходят в следующую популяцию без изменения генов;
- все «худшие» особи переходят в следующую популяцию только после применения к ним оператора мутации;
- гены из группы «другие» выбираются методом рулетки [9] и скрещиваются со случайно выбранными «лучшими» особями, а потомки, полученные после скрещивания, переходят в новую популяцию.



**Рисунок 1.** Разделение популяции на три группы.



**Рисунок 2.** Получение новой популяции.

## 8. Скрещивание

В скрещивании участвуют только особи из групп «лучшие» и «другие». Пары для формирования потомков при помощи оператора скрещивания формируются следующим образом: первый родитель выбирается случайным образом из группы «лучших», а второй родитель выбирается из числа «других» методом рулетки (то есть вероятность выбора конкретной особи из группы зависит от значения приспособленности этой особи).

После проведения отбора родителей к ним применяется оператор скрещивания. В качестве оператора скрещивания был выбран классический одноточечный кроссинговер [9] с небольшой модификацией.

Хромосомой в этой реализации генетического алгоритма является массив из чисел, отвечающих за число нейронов на конкретном скрытом слое нейронной сети. При проведении операции скрещивания возможны два разных исхода событий: длина хромосом родителей будет совпадать (одинаковое число скрытых слоёв) или будет различной.

В случае одинаковой длины хромосом применяется одноточечный кроссинговер без модификации: выбирается случайная точка разрыва, после чего потомок формируется из склейки срезов хромосом первого (от начала до точки разрыва) и второго (от точки разрыва и до конца) родителей (рисунок 3).

В случае, когда длины хромосом родителей не совпадают, нижняя часть архитектуры скрещивается по обычным правилам, а с некоторой вероятностью (в алгоритме было установлено 50%) к потомку добавляется верхняя часть архитектуры более длинной хромосомы (рисунок 4).

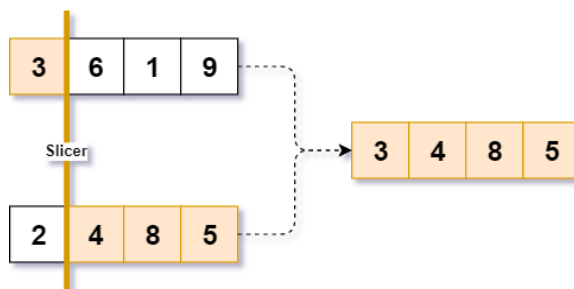


Рисунок 3. Одноточечный кроссинговер.

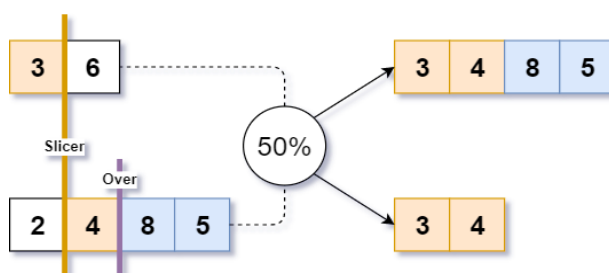


Рисунок 4. Одноточечный кроссинговер с модификацией.

## 9. Мутация

К группе «худших» особей популяции применяется оператор мутации, без возможности особей этой группы участвовать в скрещивании. При мутации каждый элемент хромосомы мутирует (заменяется на случайное число) с некоторой вероятностью, заданной в параметрах алгоритма (рисунок 5).

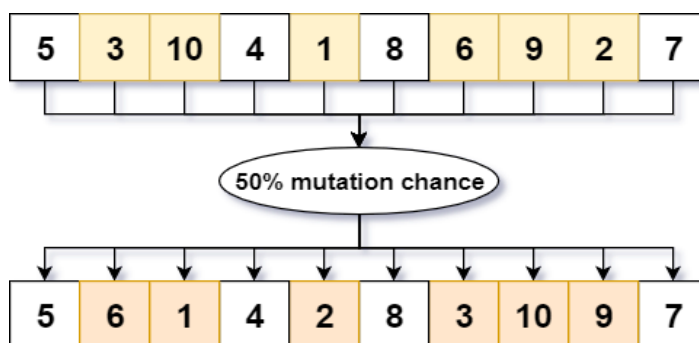


Рисунок 5. Применение оператора мутации.

## 10. Вычислительный эксперимент

Для исследования генетического подхода для настройки структуры нейронной сети был разработан модуль, реализующий искусственную нейронную сеть, а также генетический алгоритм на языке программирования Python 3.6.9.

В качестве исходной задачи, для которой выполняется настройка структуры искусственной нейронной сети, была выбрана задача диагностики неисправностей синхронного двигателя на постоянных магнитах (permanent magnet synchronous motor, PMSM).

Настраиваемые параметры генетического алгоритма приведены в таблице 1.

**Таблица 1.** Настраиваемые параметры генетического алгоритма.

№	Настраиваемый параметр	Значение настраиваемого параметра
1	Оператор селекции	Гибрид рулеточного метода и элитарной селекции
2	Оператор скрещивания	Модифицированный одноточечный кроссинговер
3	Шанс мутации конкретного гена особи	0.5

Настраиваемые параметры нейронной сети приведены в таблице 2.

Функция активации не настраивается отдельно для конкретного нейрона или слоя нейронов. Во всех примерах используется логистическая функция (сигмоида).

**Таблица 2.** Настраиваемые параметры искусственной нейронной сети.

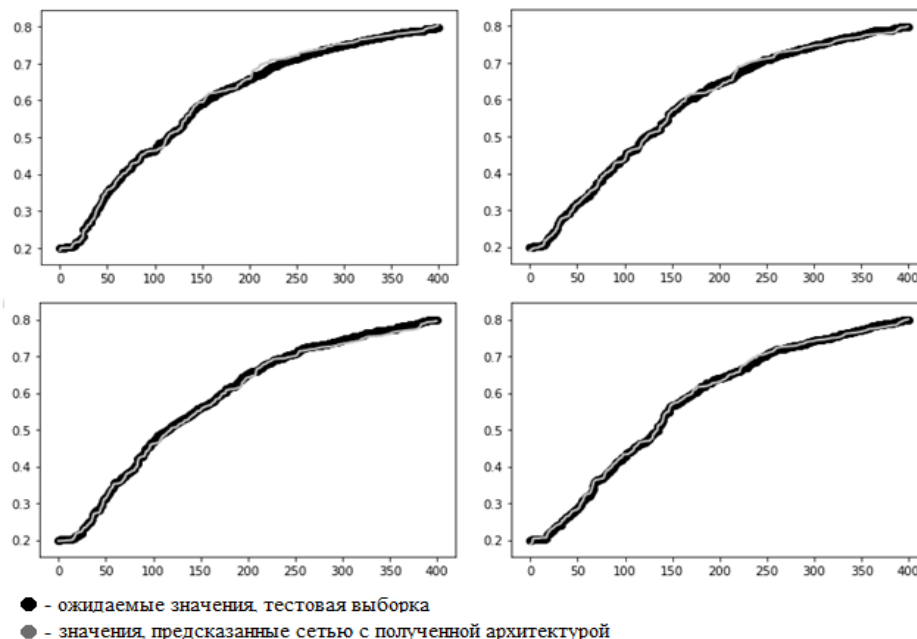
№	Настраиваемый параметр	Значение настраиваемого параметра
1	Максимальное число скрытых слоев	[1, 3]
2	Максимальное число нейронов на каждом слое	[1, 30]

Каждая нейронная сеть обучалась с использованием метода обратного распространения ошибки с коэффициентом скорости обучения  $\eta$  равным 0.01. Размер батча 200, количество эпох 50.

В результате работы генетического алгоритма с размером популяции в 100 особей и количеством эпох 50, была получена особь, показавшая наилучшие результаты предсказания нужного значения среди всех имеющихся особей на каждой эпохе. Таким образом была получена оптимальная структура нейронной сети для решения данной задачи: 11 нейронов на входном слое, 23 нейрона на первом скрытом слое, 11 нейронов на втором скрытом слое, 8 нейронов на третьем скрытом слое и 1 нейрон на выходном слое.

Результаты предсказания нейронной сети с архитектурой, полученной в ходе запуска генетического алгоритма, представлены на рисунке 6. Всего было проведено несколько запусков нейронной сети (каждый раз на новом наборе тестовых данных) с полученной архитектурой, чтобы удостовериться в качестве результата.





**Рисунок 6.** Значение температуры PMSM.

## 11. Заключение

Данная реализация генетического алгоритма показала отличные результаты в задаче подбора оптимальной архитектуры нейронной сети. Сеть с подобранной генетическим алгоритмом архитектурой показала 0.4% ошибки на тестовой выборке, что является отличным результатом для данной задачи [3], а поиск оптимальной архитектуры в этом случае не превышал 4 часов реального времени.

Выигрыш по скорости работы заметен в сравнении с полным перебором комбинаций скрытых слоёв, содержащих от 1 до 30 нейронов. По приблизительным подсчётам, такой поиск занял бы более 54 часов только на обучение сетей каждой популяции (без учёта времени на применение генетических операторов). Как было ранее отмечено в подобных исследованиях [4], полный перебор всех возможных решений занимает слишком много времени, а, следовательно, использования генетического алгоритма для решения задачи полностью оправдано.

## Список литературы

1. Arifovic, A. Using genetic algorithms to select architecture of a feedforward artificial neural network / A. Arifovic, R. Gencay // PHYSICA A: STATISTICAL MECHANICS AND ITS APPLICATIONS. – 2000. – № 289. – P. 574-594.

2. Ахметвалеев, А.М. Редукция нейросетевых моделей на основе метода двухэтапной генетической оптимизации / А.М. Ахметвалеев, А.С. Катасёв, А.П. Кирпичников // Вестник Казанского технологического университета. – 2017. – № 20-9. – С. 71-75.
3. Ефимов, С.Н. О применении параллельных генетических алгоритмов для автоматизации нейросетевого моделирования / С.Н. Ефимов, Е.С. Семенкин, В.С. Тынченко, В.В. Тынченко // Сибирский журнал науки и технологий. – 2009. – № 6-13.
4. Истомина, Ю.А. О проблематике моделирования нейронных сетей произвольной архитектуры / Ю.А. Истомина // Решетневские чтения. – 2012. – С. 609-610.
5. Истомина, Ю.А. Особенности процесса проектирования нейросетевых моделей произвольной архитектуры / Ю.А. Истомина // Актуальные проблемы авиации и космонавтики. – 2014. – С. 313-314.
6. Истомина, Ю.А. Разработка библиотеки классов генетических алгоритмов для подбора наилучшей структуры ИНС / Ю.А. Истомина, В.В. Тынченко // Актуальные проблемы авиации и космонавтики. – 2013. – С. 363-365.
7. Комарцова, Л.Г. Проектирование нейросетевых архитектур для решения прикладных задач / Л.Г. Комарцова // Известия ТРТУ. – 2005.
8. Лобин, М.А. Машинное обучение в экономике / М.А. Лобин, И.А. Филиппова // Вестник Ульяновского государственного технического университета. – 2019. – С. 68-71.
9. Панченко, Т.В. Генетические алгоритмы: учебно-методическое пособие / Т.В. Панченко; под ред. Ю. Ю. Тарасевича. – Астрахань: Издательский дом «Астраханский университет», 2007. – 87 с.
10. Сараев, П.В. Конструктивный подход построения нейроструктурных моделей на основе блочного псевдообращения / П.В. Сараев // Вестник Воронежского государственного технического университета. – 2012.
11. Соломко, Д.С. Тенденции применения нейронных сетей в юриспруденции / Д.С. Соломко, К.С. Антониади, О.И. Коваль // Colloquium-journal. – 2019.
12. Фаустова, К.И. Нейронные сети: применение сегодня и перспективы развития / К.И. Фаустова // Территория науки. – 2017. – С. 83-87.