

## СРАВНЕНИЕ АЛГОРИТМОВ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ С БИНАРНЫМИ ВХОДАМИ

© 2016 И.А. Лёзин, В.В. Муравьев

Самарский национальный исследовательский университет имени академика С.П. Королёва

Статья поступила в редакцию 11.11.2016

Данная статья сравнивает несколько алгоритмов обучения нечеткой нейронной сети Ванга-Менделя: алгоритм наискорейшего спуска, генетический алгоритм, алгоритм имитации отжига, алгоритм имитации роя частиц, алгоритм дифференциальной эволюции. Сравнение проводилось на данных из международного репозитория бинарными значениями. При решении задачи классификации наилучшие результаты показал генетический алгоритм.

*Ключевые слова:* алгоритм обучения, нейронная сеть, алгоритм наискорейшего спуска, генетический алгоритм, алгоритм имитации отжига, алгоритм имитации роя частиц, алгоритм дифференциальной эволюции.

### ПОСТАНОВКА ЗАДАЧИ

Нейронная сеть, как правило, используется тогда, когда неизвестен точный вид связей между входами и выходами, – если бы он был известен, то связь можно было бы моделировать непосредственно. Другая существенная особенность нейронных сетей состоит в том, что нейронная сеть является универсальным аппроксиматором, т.е. способна реализовывать любую функцию при должном обучении.

Для решения задачи классификации была выбрана сеть Ванга-Менделя. Сеть Ванга-Менделя относится к классу нечетких продукционных нейронных сетей, которые обладают возможностью работы с неопределенными, неполными или неточными данными.

Основной проблемой при работе с нейронной сетью является выбор способа ее обучения. Распространенным методом является метод обратного распространения ошибки, используемый вместе с градиентными алгоритмами обучения. Недостатком градиентных методов является гарантированное нахождение локального, а не глобального минимума. Стохастические методы обучения, такие как генетический алгоритм или алгоритм имитации отжига, позволяют избежать заикливания алгоритма в точке локального минимума за счет применения как детерминированного, так и случайного этапов преобразования нейронной сети [1].

*Лёзин Илья Александрович, кандидат технических наук, доцент кафедры «Информационные системы и технологии». E-mail: ilyozin@yandex.ru*

*Муравьев Вячеслав Вячеславович, студент второго курса магистратуры факультета информатики. E-mail: muravjevslava@gmail.com*

### СТРУКТУРА И ФУНКЦИОНАЛЬНЫЕ ОСОБЕННОСТИ СЕТИ

Сеть Ванга-Менделя представляет собой четырехслойную структуру, в которой первый слой выполняет фуззификацию входных переменных, второй – агрегирование значений активации условия, третий (линейный) – агрегирование  $M$  правил вывода (первый нейрон) и генерацию нормализующего сигнала (второй нейрон), тогда как состоящий из одного нейрона выходной слой осуществляет нормализацию, формируя выходной сигнал  $u(x)$ . Только первый и третий слои являются параметрическими. В первом слое это параметры функции фуззификации  $(c_j^{(k)}, \sigma_j^{(k)}, b_j^{(k)})$ , а в третьем слое – веса  $v_1, v_2, \dots, v_M$ , интерпретируемые как центр  $c_k$  функции принадлежности следствия  $k$ -го нечеткого правила вывода.

Наиболее важным фактором, влияющим на качество решения задачи с помощью нейронных сетей, является структура нейронной сети. Так как сеть Ванга-Менделя имеет фиксированное число слоев, то основной метод влияния на структуру сети – изменение количества нейронов в слоях.

От количества нейронов в сети также зависит количество операций на одной итерации обучения. С одной стороны, уменьшение количества нейронов ведет к сокращению времени обучения сети. С другой стороны, малое количество нейронов ведет к увеличению ошибки обучения [2].

Также стоит учитывать функциональные особенности сети. Сеть Ванга-Менделя построена на основе нечеткой модели Мамдани-Заде. В модели Мамдани-Заде определены 3 вида операций – Т-норма ( $\otimes$ ), S-норма ( $\oplus$ ) и импликация ( $\rightarrow$ ). Совокупность данных операций называется алгеброй нечеткого вывода. В данной работе рассматривались следующие алгебры, предложенные в работе [1]:

Алгебра Гёделя

$$\mu_A(x) \otimes \mu_B(y) = \min\{\mu_A(x), \mu_B(y)\}$$

$$\mu_{A \rightarrow B}(x, y) = \begin{cases} 1, \mu_A(x) \leq \mu_B(y) \\ \mu_B(y), \mu_A(x) > \mu_B(y) \end{cases}$$

$$\mu_A(x) \oplus \mu_B(y) = \max\{\mu_A(x), \mu_B(y)\}$$

Алгебра Гогена

$$\mu_A(x) \otimes \mu_B(y) = \mu_A(x)\mu_B(y)$$

$$\mu_{A \rightarrow B}(x, y) = \begin{cases} 1, \mu_A(x) \leq \mu_B(y) \\ \mu_B(y) / \mu_A(x), \mu_A(x) > \mu_B(y) \end{cases}$$

$$\mu_A(x) \oplus \mu_B(y) = \mu_A(x) + \mu_B(y) - \mu_A(x)\mu_B(y)$$

Алгебра Лукашевича

$$\mu_A(x) \otimes \mu_B(y) = \max\{0, \mu_A(x) + \mu_B(y) - 1\}$$

$$\mu_{A \rightarrow B}(x, y) = \min\{1, 1 - \mu_A(x) + \mu_B(y)\}$$

$$\mu_A(x) \oplus \mu_B(y) = \min\{1, \mu_A(x) + \mu_B(y)\}$$

### АЛГОРИТМЫ ОБУЧЕНИЯ СЕТИ

Также важным фактором, влияющим на качество решения задачи, является выбор алгоритма обучения. В рамках данной работы рассматривалось обучение сети при помощи следующих алгоритмов:

Градиентные алгоритмы:

1. Алгоритм наискорейшего спуска и метод обратного распространения ошибки (ОРО)[2]

Стохастические алгоритмы:

1. Генетический алгоритм (ГА) [3]

2. Алгоритм имитации отжига (АИО) [2]

3. Алгоритм имитации роя частиц (АИРЧ) [4]

4. Алгоритм дифференциальной эволюции (АДЭ) [5]

### АВТОМАТИЗИРОВАННАЯ СИСТЕМА

Экспериментальные исследования производились на специально разработанной системе для обучения и тестирования нейросетевых классификаторов на основе описанных алгоритмов. Система обладает следующими функциями:

- выбор обучающей и тестовой последовательностей для сети;
- ввод и редактирование параметров нейронной сети, алгоритма обучения;
- обучение нейронной сети;
- визуализация результатов обучения и тестирования сети.

### ЭКСПЕРИМЕНТАЛЬНЫЕ ДАННЫЕ И ПРЕПРОЦЕССИНГ

В качестве данных тестовых задач были взяты задачи классификации элементов интернет-страниц на рекламные и нерекламные из репозитория UCІ Калифорнийского университета [6].

Выбранные данные состоят из множества булевых атрибутов, объединенных авторами в несколько «надклассов». В целях оптимизации процесса обучения сетей данные прошли пре-процессинг аддитивным алгоритмом.

Пусть дано  $C = \{C_1, C_2, \dots, C_n\}$  - множество надклассов атрибутов логических данных. В обработанных данных сопоставим надклассу  $C_i$  значение  $S_i$  - количество «истинных» (равных 1) значений в надклассе  $C_i$ .

Данные были разделены на обучающую и тестовую выборки в соотношении 20:80 соответственно.

### ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ И ИХ ОЦЕНКА

Как показатель качества обучения для количественной оценки разности между ожидаемым и реальным значениями на выходах нейронной сети рассчитывались суммарное среднеквадратическое отклонение (СКО) обучения, определяемое формулой

$$СКО = \sqrt{\frac{1}{V-1} \cdot \frac{1}{E} \cdot \sum_{t=1}^V \sum_{j=1}^E (y_{j,t} - d_{j,t})^2},$$

а также погрешность классификации, равная отношению количества неверно распознанных данных к общему размеру выборки.

Выбранные оценки качества будем рассчитывать как на этапе обучения, так и на этапе тестирования нейронных сетей. В дальнейшем будем называть их погрешностями обучения и тестирования соответственно. Погрешность обучения показывает, насколько хорошо нейронная сеть способна накапливать информацию об обучающих выборках. Погрешность тестирования характеризует способность сети к обобщению полученных знаний, то есть находить зависимость выходных сигналов от входных данных, не участвующих в обучении.

### РЕЗУЛЬТАТЫ

В ходе исследований было выяснено, что ОРО, ГА, АИО, АИРЧ и АДЭ качественно обучают сеть нечеткую сеть Ванга-Менделя. Хотя на выбранных данных ОРО и АДЭ показали результаты хуже, чем остальные.

Относительное сравнение результатов различных алгоритмов показало:

ОРО – в основном требует большого количества итераций обучения – от 5000 и более (ГА зачастую хватает и 1000 итераций). Также дает погрешность классификации 0.095–0.1, что хуже ГА, АИО и АИРЧ.

ГА требует большее количество нейронов по сравнению с применением остальных алгоритмов. С другой стороны, ГА обучает сети, классифи-

**Таблица 1.** СКО обучения и погрешность классификации при решении задач классификации элементов интернет-страниц

Тип алгоритма и алгебра	СКО обучения	Погрешность классификации	Количество нейронов	Количество итераций
ОРО, Гёделя	0,351	0,105	27	5000
ОРО, Гогена	0,331	0,103	27	5000
ОРО, Лукашевича	0,334	0,096	27	1000
ГА, Гёделя	0,259	0,067	81	5000
ГА, Гогена	0,270	0,068	45	1000
ГА, Лукашевича	0,258	0,064	45	1000
АИО, Гёделя	0,292	0,086	81	5000
АИО, Гогена	0,298	0,093	27	5000
АИО, Лукашевича	0,324	0,095	27	5000
АИРЧ, Гёделя	0,277	0,077	27	50
АИРЧ, Гогена	0,250	0,073	27	50
АИРЧ, Лукашевича	0,346	0,075	27	50
АДЭ, Гёделя	0,304	0,085	27	500
АДЭ, Гогена	0,318	0,109	45	500
АДЭ, Лукашевича	0,295	0,091	27	100

**Таблица 2.** Дополнительные параметры алгоритмов, при которых получены оптимальные значения СКО и погрешности классификации

Параметр алгоритма/ Алгебра	Гёделя	Гогена	Лукашевича
ОРО, Шаг обучения	0,05	0,01	0,1
ГА, Размер популяции	30	30	50
ГА, Коэффициент кроссовера	0,5	0,5	0,5
АИО, Начальная температура	0,001	0,05	0,001
АИРЧ, Пар-р $\omega$	0,8	0,5	1,2
АИРЧ, Пар-р $\varphi_p$	0,5	0,8	0,5
АИРЧ, Пар-р $\varphi_g$	1,2	1,2	1,2
АДЭ, Вероятность изменения	0,25	0,25	0,5
АДЭ, Коэффициент кроссовера	1,5	2	0,5

цирующие с погрешностью менее 0.07 – лучший результат среди всех алгоритмов.

АИО показывает средние результаты обучения (0.8-0.9), требуя среднее количество временных ресурсов для обучения.

АИРЧ может обучать сеть только в «оффлайн» режиме, в отличие от остальных алгоритмов, к

тому же требует метаоптимизация его параметров требует много больше ресурсов, чем для других алгоритмов.

Хотя АДЭ схож с ГА по общей концепции, но показывает результаты хуже (0.085 и 0.064 соответственно). С другой стороны, АДЭ требуется меньше итераций обучения (500 и 1000 соответственно).

## СПИСОК ЛИТЕРАТУРЫ

1. Лёзин И.А., Солдатова О.П. Исследование качества решения задачи классификации нейронными нечёткими продукционными сетями на основе модели вывода Мамдани–Заде // Вестн. Сам. гос. техн. ун-та. Сер. Физ.-мат. науки, 2014. 2(35). С. 136–148.
2. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002. 344 с.
3. Лёзин И.А., Муравьев В.В. Применение генетического алгоритма для обучения нечеткой многовыходовой нейронной сети Ванга-Менделя // Перспективные информационные технологии (ПИТ 2014): труды Международной научно-технической конференции [под ред. С.А. Прохорова]. Самара: Издательство Самарского научного центра РАН, 2014. С.64 – 66.
4. Метод роя частиц [Электронный ресурс]. URL: [https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4\\_%D1%80%D0%BE%D1%8F\\_%D1%87%D0%B0%D1%81%D1%82%D0%B8%D1%86](https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D1%80%D0%BE%D1%8F_%D1%87%D0%B0%D1%81%D1%82%D0%B8%D1%86) (дата обращения 10.09.2016).
5. Дифференциальная эволюция [Электронный ресурс]. URL: [https://ru.wikipedia.org/wiki/%D0%94%D0%B8%D1%84%D1%84%D0%B5%D1%80%D0%B5%D0%BD%D1%86%D0%B8%D0%B0%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F\\_%D1%8D%D0%B2%D0%BE%D0%BB%D1%8E%D1%86%D0%B8%D1%8F](https://ru.wikipedia.org/wiki/%D0%94%D0%B8%D1%84%D1%84%D0%B5%D1%80%D0%B5%D0%BD%D1%86%D0%B8%D0%B0%D0%BB%D1%8C%D0%BD%D0%B0%D1%8F_%D1%8D%D0%B2%D0%BE%D0%BB%D1%8E%D1%86%D0%B8%D1%8F) (дата обращения 10.09.2016).
6. UCI Machine Learning Repository [Электронный ресурс]. URL: <http://archive.ics.uci.edu/ml/> (дата обращения 10.09.2016).

## COMPARISON OF LEARNING ALGORITHMS FOR NEURAL NETWORKS WITH BINARY INPUTS

©2016 I.A. Lyozin, V.V. Muravyov

Samara National Research University named after Academician S.P. Korolyov

The article compares several learning algorithms for Wang-Mendel's fuzzy neural network: steepest descent algorithm, genetic algorithm, simulated annealing, simulate particle swarm algorithm, differential evolution algorithm. The comparison has been processed on the data from the international repository with binary values. The genetic algorithm shows best results for solving the classification problem.

*Keywords:* learning algorithm, neural network, steepest descent algorithm, genetic algorithm, simulated annealing, simulate particle swarm algorithm, differential evolution algorithm.

*Ilya Lyozin, Candidate of Technics, Associate Professor at the Information Systems and Technologies Department.*

*E-mail: ilyozin@yandex.ru.*

*Vyacheslav Muravyov, Second-Year Student of Magistracy of the Faculty of Computer Science.*

*E-mail: muravjevslava@gmail.com.*