

Сравнительный анализ алгоритмов обучения искусственной нейронной сети

Е.В. Пучков

Методы обучения искусственных нейронных сетей (ИНС) разделяют на четыре категории: нулевого порядка, первого порядка, второго порядка и нелокальные модификации одноточечных методов. Поскольку целевая функция в задачах обучения многоэкстремальная, для нахождения глобального экстремума используют генетические алгоритмы, методы облака, рестартов, поколений, модифицированный метод многогранника. Более подробно с ними можно ознакомиться в [1 - 4].

В основе многих алгоритмов обучения ИНС типа многослойный персептрон лежит итерационный процесс корректировки весовых коэффициентов w , для которого необходимо выбрать направление движения и шаг [1]:

$$w(t + 1) = w(t) + \eta(t)p(t), \quad (1)$$

где вектор $p(t)$ – направление движения; $\eta(t)$ – размер шага; t – порядковый номер текущей итерации. Благодаря известному методу обратного распространения ошибки можно произвести декомпозицию сложной задачи обучения.

В статье проводится сравнение алгоритмов обучения ИНС таких, как генетический [5], адаптивный [6] и гибридный [7]. Используются данные для задачи классификации «Ирисы Фишера» [8], в качестве экспериментальной среды выбрана веб-версия нейроэмулятора *NeuroNADS* [9, 10]. Построены комитеты размером 10 полносвязных двухслойных нейронных сетей с применением каждого алгоритма. Количество нейронов в скрытом слое – 9 (согласно теореме Колмогорова-Арнольда-Хехт-Нильсена). Критерий остановки обучения ИНС – среднеквадратическая ошибка со значением 0,01, размер шага $\eta(t)$ – 0,1. Примеры из обучающей выборки подавались случайно. В качестве функции активации нейронов

использовались простые сигмоиды.

В последнее время часто решают задачи поиска глобального экстремума с помощью генетических алгоритмов. Обучение ИНС с помощью генетического алгоритма происходит в два этапа. На первом – формируются новые хромосомы, на втором – отбираются наилучшие хромосомы в популяцию. В качестве хромосомы выступают веса нейронов ИНС.

$$\vec{X}_i = \{w_{11}^{i_1}, w_{12}^{i_1}, \dots, w_{1a}^{i_1}, w_{21}^{i_1}, \dots, w_{aa}^{i_1}, w_{11}^{i_2}, \dots, w_{1b}^{i_3}\},$$

где X – хромосома; i – индекс представителя популяции; w – вес нейрона.

Новые хромосомы образуются в результате скрещивания и различных видов мутаций, которые можно проводить в произвольном порядке в пределах одной эпохи. Хромосома, соответствующая заданному критерию качества (в случае с ИНС – это ошибка обучения), будет отобрана в качестве наилучшей.

Результаты применения генетического алгоритма представлены на рис.1.

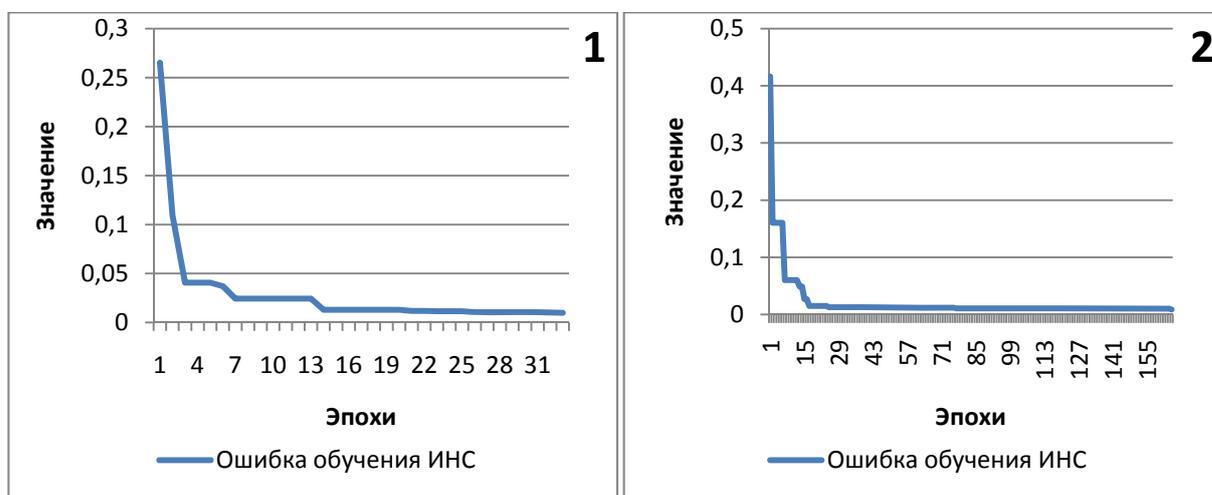


Рис. 1. – График зависимости ошибки обучения от количества эпох для генетического алгоритма (1–лучший, 2–худший результаты)

Генетический алгоритм справился с поставленной задачей, но интенсивность изменения ошибки обучения резко снижалась после 15 эпох. Среднее время обучения ИНС составило 109 эпох.

В методе сопряженных градиентов, который относится к методам обучения первого порядка, формулу поиска минимума на основе предыдущих направлений можно обобщить [5]:

$$p(t + 1) = g(t) + \sum_{i=1}^{\min(t-1,n)} \beta(i)g(t - i), \quad (2)$$

где вектор $p(t)$ – направление движения; $g(t)$ – направление антиградиента на текущей итерации t ; $\beta(i)$ – коэффициент, определяющий вес i -го градиента; n – количество запоминаемых градиентов. При $n = 0$ получим простой градиентный спуск, а при $n = \infty$, суммируя все предыдущие направления – методы сопряженных градиентов. Настройка параметра n и последовательностей $\eta(t)$, $\beta(i)$ из формул (1) и (2), соответственно, позволит использовать более гибкое решение – адаптивный алгоритм обучения ИНС [5].

На рис. 2 представлены графики зависимости ошибки обучения от количества эпох для адаптивного алгоритма.

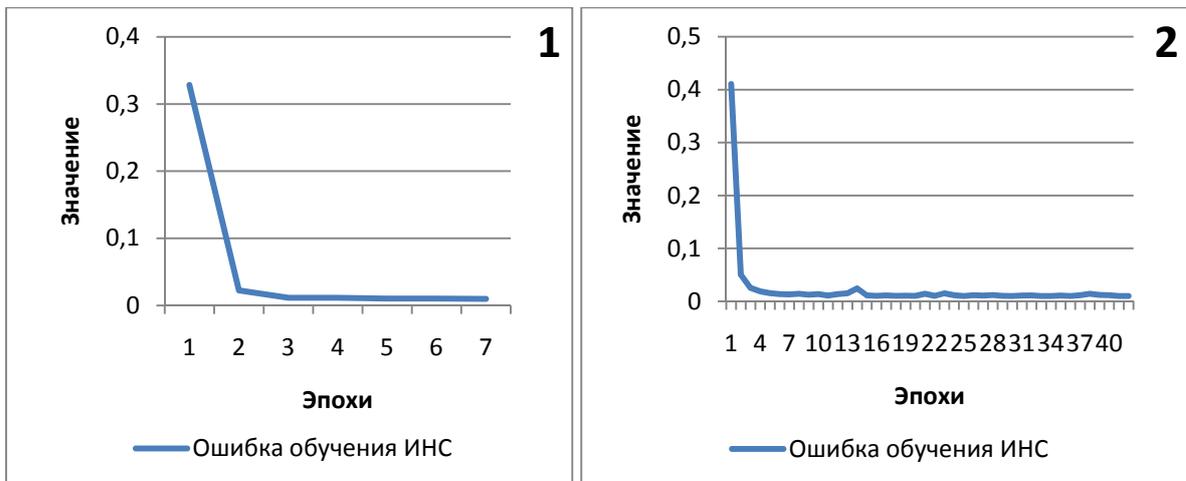


Рис. 2. – График зависимости ошибки обучения от количества эпох для адаптивного алгоритма (1–лучший, 2–худший результаты)

Среднее время обучения ИНС с применением адаптивного алгоритма составило 22 эпохи, что в 5 раз меньше, чем среднее время обучения генетическим алгоритмом. Сходимость алгоритма к локальному минимуму наступала за 5-7 эпох.

Последний метод, который мы будем использовать для обучения ИНС, основан на последовательном применении адаптивного и генетического

алгоритмов [6]. Отметим, что при переходе к генетическому алгоритму добавляется к популяции хромосома – ИНС, обученная адаптивным алгоритмом. В качестве критерия перехода используется значение ошибки обучения равное 0,015. На рис. 3 изображены графики зависимости ошибки обучения от количества эпох для гибридного метода.

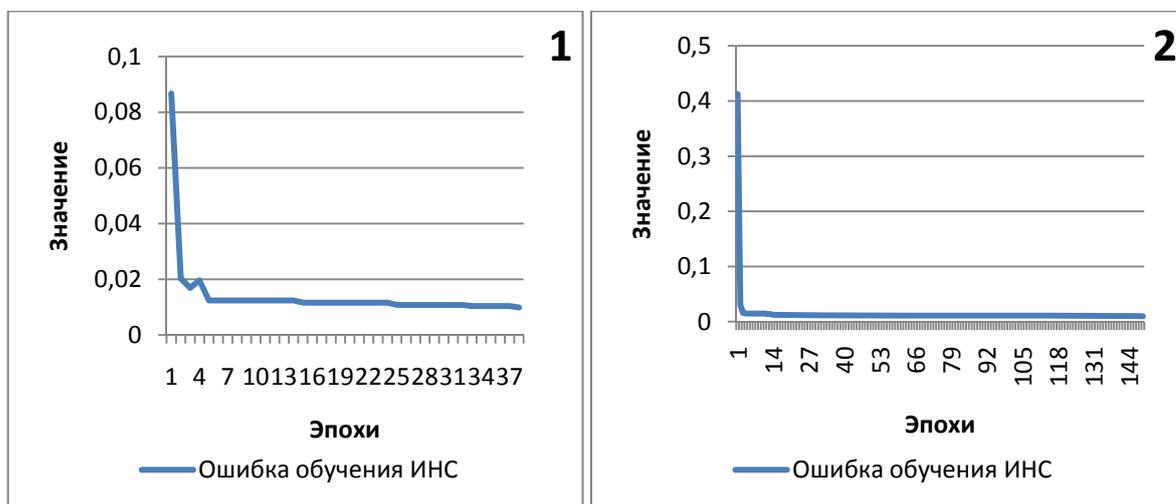


Рис. 3. – График зависимости ошибки обучения от количества эпох для гибридного метода (1–лучший, 2–худший результаты)

По результатам сравнения адаптивный алгоритм сходится быстрее, чем генетический и гибридный (таблица №1). Окончательный выбор алгоритма будет зависеть от конкретный задачи, потому что в задаче со сложной целевой многоэкстремальной функцией, в которой с помощью гибридного алгоритма можно быстрее вычислить решение в окрестности глобального минимума, адаптивный алгоритм может показать результаты хуже [7].

Отметим, что данные результаты не могут в полной мере определять точность классификации и точность метода, поскольку не проводилась оценка количества распознанных экземпляров ириса и не исследовалась обобщающая способность построенных ИНС.

Таблица №1. Время обучения ИНС (эпохи)

Номер ИНС	Алгоритм		Гибридный метод
	генетический	адаптивный	
1	32	20	38
2	54	21	86
3	70	16	71
4	130	13	35
5	101	35	96
6	156	20	131
7	80	12	52
8	168	41	148
9	136	38	37
10	164	6	103
Среднее	109	22	80

Литература

1. Хайкин С. Нейронные сети: полный курс, 2-е изд.; – пер. с англ./ С. Хайкин. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
2. Тархов Д.А. Нейронные сети. Модели и алгоритмы. Кн.18: справочное издание. (Серия "Нейрокомпьютеры и их применение"). – М. : Радиотехника, 2005. – 256 с.
3. Бодянский Е.В., Руденко О.Г. Искусственные нейронные сети: архитектуры, обучение, применения. – Харьков: ТЕЛЕТЕХ, 2004. – 369 с.
4. Осовский С. Нейронные сети для обработки информации / Пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.
5. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы; пер. с польск. И. Д. Рудинского. – М.: Горячая линия -Телеком, 2006. – 452 с.
6. Лиля В.Б. Алгоритм и программная реализация адаптивного метода обучения искусственных нейронных сетей [Электронный ресурс] // Инженерный вестник Дона, 2012, №1. – Режим доступа: <http://ivdon.ru/magazine/archive/n1y2012/626> (доступ свободный) – Загл. с экрана. – Яз. рус.

7. Бежавский Г.И., Пучков Е.В., Лиля В.Б. Алгоритм и программная реализация гибридного метода обучения искусственных нейронных сетей // Программные продукты и системы. – Тверь, 2012. – №4. – с. 96 - 100.
8. Iris Data Set. The UCI Machine Learning Repository [Электронный ресурс]. – Режим доступа: <http://archive.ics.uci.edu/ml/datasets/Iris> (доступ свободный) – Загл. с экрана. – Яз. англ.
9. Пучков Е.В. Применение нейроэмулятора «NeuroNADS» для определения ступени и времени торможения при управлении горочными замедлителями [Электронный ресурс] // Инженерный вестник Дона, 2010, №4. – Режим доступа: <http://ivdon.ru/magazine/latest/n4y2010/263> (доступ свободный) – Загл. с экрана. – Яз. рус.
10. NeuroNADS [Электронный ресурс]: веб-сервис — Режим доступа: <http://www.service.i-intellect.ru/emulator.php> (ограниченный доступ) – Загл. с экрана. – Яз. рус.