УДК 519.68

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ МОДИФИЦИРОВАННОГО АЛГОРИТМА ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ

Матвеева Е.А.,
научный руководитель канд. техн. наук Липинский Л.В.
Сибирский государственный аэрокосмический университет им. М. Ф. Решетнева

Исследуется эффективность модификации алгоритма обучения нейронных сетей. Разработана программная система, реализующая алгоритм. Получены результаты на тестовых и практических задачах.

Искусственные нейронные сети, в наши дни, успешно применяются для аппроксимации, классификации, прогнозирования и т.д. В задачах, где вычисление целевой функции связано с материальными или временными затратами (производственные процессы, медицинские исследования, физические и химические опыты и т.п.) возникает вопрос о сокращении вычислений целевой функции при нейросетевом моделировании. В работе предлагается предварительно обучающую выборку отсортировать таким образом, чтобы на нейронную сеть подавались обучающие пары максимально отдаленные друг от друга. Для включения точки в обучающую выборку определяются минимальные расстояние между точкой из оставшейся части выборки и всеми точками из обучающей выборки. В обучающую выборку включается та точка, чье минимальное расстояние до точек из обучающей выборки, максимально. Такая сортировка должна позволить достигать желаемой ошибки обучения при меньшей обучающей выборке. Также было предложено равномерно увеличивать объем выборки и использовать все данные только на последних итерациях обучения.

Введем понятия:

- Упорядоченная выборка - выборка, упорядоченная в соответствии с предложенной выше модификацией;
- Неупорядоченная выборка – первичная выборка.

Тестирование осуществлялось стократным прогоном каждого алгоритма. Результаты сведены в таблицу 1. Ошибка обучения – ошибка на обучающей выборке. Ошибка обобщения – ошибка на тестовой выборке. Число вычислений – число вычислений нейросетевой модели. Используется для оценки скорости обучения алгоритма. Под скоростью будем понимать количество вычислений значений целевой функции, которое равняется количеству эпох, требуемых для достижения желаемой ошибки умноженное на объём выборки.

Результаты исследования сведены в таблицу 1:

Таблица 1. Результаты исследования на функции $f(x) = x^2$

Объём выборки, %	**Выборка**	**Ошибка обучения**	**Ошибка обобщения**	**Число вычислений**
15	Упорядоченная	0,0074651517306	0,0099756865431	30855
	Неупорядоченная	0,0075655369729	0,0099911976149	34155
30	Упорядоченная	0,0021736375576	0,0099975296132	142825
	Неупорядоченная	0,0020870184198	0,0099922108124	159268
50	Упорядоченная	0,0082708511497	0,0099949340895	77679
	Неупорядоченная	0,0247694709496	0,0099839510001	221646

80	Упорядоченная	0,0126484457314	0,0099439032858	81686
	Неупорядоченная	0,0205352489359	0,0099453250457	91008
100	Упорядоченная	0,0080763055653	0,0099881486321	116679
	Неупорядоченная	0,0115387575205	0,0099775855565	148986
Меняется от 10 до 80	Упорядоченная	0,0057034572809	0,0170345723055	18900
	Неупорядоченная	0,0087034570949	0,0340875364089	18900
Меняется от 20 до 100	Упорядоченная	0,0020870184198	0,0090560836102	27300
	Неупорядоченная	0,0383951018414	0,0115383951000	27300

Таблица 2. Результаты исследования на функции $f(x) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$

Объём выборки, %	**Выборка**	**Ошибка обучения**	**Ошибка обобщения**	**Число вычислений**
15	Упорядоченная	0,025797693524	0,034990976017	36270
	Неупорядоченная	0,026928677322	0,034995744105	66040
30	Упорядоченная	0,022505598345	0,034981812545	59384
	Неупорядоченная	0,022392652396	0,034971413097	87804
50	Упорядоченная	0,018792955542	0,034998534654	67645
	Неупорядоченная	0,016314603944	0,034887901345	119973
80	Упорядоченная	0,019603536541	0,034959295264	96640
	Неупорядоченная	0,019136595928	0,034998251700	626952
100	Упорядоченная	0,022133333113	0,034994599247	468830
	Неупорядоченная	0,021328828047	0,034983708804	627125
Меняется от 10 до 80	Упорядоченная	0,034990976017	0,050187927231	18100
	Неупорядоченная	0,067499059592	0,055787925171	18100
Меняется от 20 до 100	Упорядоченная	0,013339265265	0,03488767889	22500
	Неупорядоченная	0,034990394477	0,05018792517	22500

Проанализировав данную таблицу можно сделать следующие выводы:

1. В большинстве случае сортировка обучающей выборке позволяет снизить время обучения. В некоторых случаях время обучения на упорядоченной выборке в три раза меньше, чем на исходных данных.

2. Обучение на выборке, у которой объём меняется от 20% до 100%, дает минимальное число вычислений, и минимальную ошибку обобщения.

Предложенный подход был апробирован на реальных задачах.
Из всей обучающей выборки выберем случайным образом 20% точек для экзаменующей выборки, по которой и будет находиться ошибка обобщения. Используем для обучения равномерно увеличивающийся объем выборки от 20 до 100 процентов, т.к. в исследовании, описанном выше - именно такой подход дал наилучший результат.

В качестве тестовых задач для проверки работоспособности и производительности алгоритма использовались задачи из репозитория mldb, данные которых подвергались минимальной обработке. Нечисловые данные заменялись

целочисленными классами. Большинство задач, имеющихся в репозитории, считаются сложными для решения классическими алгоритмами, к некоторым приводится информация о попытках решения другими методами.

Задача выявления заболевания печени

В базе данных присутствует 345 описаний диагностики заболеваний печени. Каждый случай имеет в описании 6 параметров.

Вывод в задаче бинарный – есть ли у больного заболевание печени или нет. Среди параметров 5 целочисленных с различными диапазонами и один вещественный. Результаты представлены в таблице 2.

Таблица 3. Результаты решения задачи выявления заболевания печени

Неправильно классифицировано	**Ошибка классификации**
6 из 35	17%

Задача классификации ирисов

В базе данных присутствует описание 149 растений ириса. 4 входных параметра: длина и ширина чашелистика и лепестка, все представлены вещественными числами. 3 класса вывода. Ошибку классификации можно увидеть в таблице ниже.

Таблица 4. Результаты классификации ирисов

Неправильно классифицировано	**Ошибка классификации**
1 из 110	0,9%

Выводы: Данные исследования привели к положительным результатам – если обучающее множество велико, то можно в ходе подбора архитектуры сети использовать лишь часть выборки. Также было выяснено, что если подавать на выборку образ каждый раз максимально удаленный от уже предъявленных НС на данной итерации, то желаемая ошибка обобщения будет достигнута быстрее, чем, если подавать образы в произвольном порядке.

Исследование данного алгоритмического обеспечения доказало его работоспособность и эффективность на тестовых функциях и ряде практических задач.