

ВЛИЯНИЕ ВИДА ФУНКЦИЙ АКТИВАЦИИ НЕЙРОНОВ НА ОТНОСИТЕЛЬНУЮ ОШИБКУ ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ

Е. А. Дегтярев, А. Л. Карякин

Оценивается влияние вида функций активации нейронов искусственных нейронных сетей на относительную ошибку прогнозирования электропотребления промышленного предприятия.

Ключевые слова: горное предприятие; искусственные нейронные сети; прогнозирование электропотребления; функции активации нейронов; цех агломерации; электрическая энергия.

Использование методов искусственного интеллекта для решения задач прогнозирования временных рядов находит практическое применение в современной электроэнергетике. Применение нейросетевых моделей в программах для составления заявок на приобретение электроэнергии позволяет существенно снизить погрешность между заявленными и потребленными объемами и соответственно затраты предприятия на штрафы при их значительном рассогласовании.

В теории нейронных сетей [1–4] не дается рекомендаций по выбору функций активации нейронов применительно к различным структурам сетей, поэтому в данной работе ставилась задача рассмотреть возможные варианты архитектуры нейронных сетей с различными функциями активации, проанализировать полученные в ходе экспериментов данные, обобщить полученный результат и выявить общие закономерности.

Экспериментирование проводилось на данных, полученных от ОАО «Качканарский ГОК» для цеха агломерации [5]. Предварительно весь совокупный объем факторов, оказывающих влияние на потребление электроэнергии, был изучен с помощью корреляционного анализа, на основе которого были выявлены наиболее значимые переменные.

Для обучения сети и экспериментов с ее архитектурой сформировано обучающее и контрольное множества, из состава которых на основании экспериментов с обобщенно-регрессионными сетями с применением метода визуального анализа диаграммы рассеяния были исключены выбросы. С целью

проведения сравнительного анализа и выявления общих закономерностей в результатах экспериментов с различными структурами сетей из исходных данных сформированы три базы. Каждая из баз содержит по 14, 15 и 16 переменных: производство агломерата за текущую смену, время работы двух агломашин, среднесуточные и крайние значения температуры уличного воздуха (наличием или отсутствием данных предикторов отличаются между собой базы данных), порядковый номер дня в году, значения потребления электрической энергии в предыдущие десять смен. Далее массивы исходных данных были разбиты на обучающее (3926 изм.), тестовое (184 изм.) и контрольное (185 изм.) множества. Для объективной оценки точности прогнозирования из состава последнего множества значения выбросов не были исключены.

Элементы теории вероятности и математической статистики позволяют обратиться к правилу трех сигм [6], из которого следует: вероятность того, что абсолютная величина отклонения превысит утроенное среднее квадратическое отклонение, очень мала, а именно равна 0,0027 % или 0,27 %. Поэтому сравнение сетей производилось по величинам математического ожидания и утроенного стандартного отклонения.

В результатах экспериментов определен верхний предел изменения относительной ошибки прогноза на тестовом множестве (предельная относительная ошибка прогноза на тестовом множестве), т. е. $\bar{x} + 3\sigma$, где \bar{x} – среднее значение относительной ошибки прогноза на тестовом множестве, определяемое формулой, отн. ед.:

$$\bar{x} = \left(\sum_{i=1}^N x_i \right) / N,$$

где x_i – значение относительной ошибки прогноза на тестовом множестве для i -го опыта, отн. ед.; N – общее число опытов (наблюдений).

Стандартное выборочное отклонение рассчитывалось в соответствии со следующим выражением, отн. ед. [7]:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}.$$

Были построены графики зависимости $\bar{x} + 3\sigma$ от вида функции активации и номера опыта, так как в первую очередь авторов интересует максимально возможная ошибка прогноза, а не то, какова доля величин, принимающих участие в ее формировании. Для определения средней относительной ошибки прогноза на тестовом множестве во всех экспериментах проводилось 30 опытов.

При синтезе сетей для каждого типа архитектуры руководствовались одними и теми же принципами, обеспечивая тем самым одинаковый порядок формирования структур сетей: число нейронов во входном слое сети

Влияние вида функций активации нейронов на величину прогноза для трехслойной сети с 16 входными переменными, 16 нейронами во входном слое и 18 нейронами в скрытом слое

Номер опыта	Вид функции активации для нейронов			\bar{x} , отн. ед.	σ , отн. ед.	$\bar{x} \pm 3\sigma$, отн. ед.	$\bar{x} + 3\sigma$, отн. ед.	Кодировка функций активаций в системах счисления	
	входного слоя	скрытого слоя	выходного слоя					двоичной	десятичной (x2)
1	logsig	purelin	logsig	0,0422	0,0021	0,0422 ± 0,0063	0,0485	011001	50
2	logsig	purelin	purelin	0,0423	0,0025	0,0423 ± 0,0075	0,0498	101001	82
3	logsig	logsig	logsig	0,0429	0,0023	0,0429 ± 0,0069	0,0498	010101	42
4	purelin	purelin	purelin	0,0490	0,0004	0,0490 ± 0,0012	0,0502	101010	84
5	logsig	tansig	purelin	0,0438	0,0022	0,0438 ± 0,0066	0,0504	100001	66
6	purelin	tansig	purelin	0,0428	0,0026	0,0428 ± 0,0078	0,0506	100010	68
7	purelin	logsig	logsig	0,0425	0,0028	0,0425 ± 0,0084	0,0509	010110	44
8	logsig	logsig	purelin	0,0435	0,0025	0,0435 ± 0,0075	0,0510	100101	74
9	tansig	purelin	purelin	0,0435	0,0025	0,0435 ± 0,0075	0,0510	101000	80
10	purelin	logsig	purelin	0,0428	0,0031	0,0428 ± 0,0093	0,0521	100110	76
11	tansig	logsig	purelin	0,0436	0,0029	0,0436 ± 0,0087	0,0523	100100	72
12	tansig	tansig	purelin	0,0446	0,0026	0,0446 ± 0,0078	0,0524	100000	64
13	tansig	logsig	logsig	0,0451	0,0054	0,0451 ± 0,0162	0,0613	010100	40
14	logsig	tansig	tansig	0,0453	0,0075	0,0453 ± 0,0225	0,0678	000001	2
15	tansig	tansig	logsig	0,0700	0,1223	0,0700 ± 0,3669	0,4369	010000	32
16	logsig	logsig	tansig	0,0679	0,1342	0,0679 ± 0,4026	0,4705	000101	10
17	logsig	tansig	logsig	0,0832	0,1569	0,0832 ± 0,4707	0,5539	010001	34
18	tansig	logsig	tansig	0,0771	0,1754	0,0771 ± 0,5262	0,6033	000100	8
19	tansig	purelin	logsig	0,0884	0,1751	0,0884 ± 0,5253	0,6137	011000	48
20	tansig	tansig	tansig	0,0792	0,1919	0,0792 ± 0,5757	0,6549	000000	0
21	purelin	tansig	logsig	0,1083	0,2036	0,1083 ± 0,6108	0,7191	010010	36
22	purelin	purelin	logsig	0,1201	0,2277	0,1201 ± 0,6831	0,8032	011010	52
23	purelin	logsig	tansig	0,1517	0,2599	0,1517 ± 0,7797	0,9314	000110	12
24	purelin	tansig	tansig	0,2073	0,3237	0,2073 ± 0,9711	1,1784	000010	4
25	purelin	purelin	tansig	0,1766	0,3676	0,1766 ± 1,1028	1,2794	001010	20
26	tansig	purelin	tansig	0,1650	0,4212	0,1650 ± 1,2636	1,4286	001000	16
27	logsig	purelin	tansig	0,3336	0,6876	0,3336 ± 2,0628	2,3964	001001	18

принимается равным числу входных переменных, число нейронов в промежуточном слое сети определяется в соответствии с формулой [2]:

$$\frac{mN}{1 + \log_2 N} \leq L_w \leq (N + m)(n + m + 1) + m,$$

где m – размерность выходного сигнала, ед.; N – число элементов обучающей выборки, шт.; n – размерность входного сигнала, ед.; L_w – необходимое число синаптических весов.

Оценив необходимое число весов, можно рассчитать число нейронов в скрытых слоях:

$$L = \frac{L_w}{n + m}.$$

Были рассмотрены следующие функции активации нейронов: *logsig* – логарифмическая (сигмоидальная); *purelin* – линейная; *tansig* – гиперболический тангенс (сигмоидальная).

По полученным в результате опытов таблицам можно оценить влияние вида функций активации нейрона на величину прогноза для однослойных сетей (\bar{x} , σ , $\bar{x} \pm 3\sigma$, $\bar{x} + 3\sigma$). Число возможных вариантов сетей для одного типа архитектуры (числа слоев) в зависимости от вида используемых функций активации нейронов определялось с помощью элементов комбинаторики.

Для того чтобы дать количественную оценку полученным результатам для трехслойных сетей, в таблицах приведены шифры вида функций активации нейронов по слоям в двоичной и десятичной системах. Кодировка в двоичной системе представлена ниже. Для наглядности результатов в таблице соответствующему двоичному коду приводится удвоенное десятичное число. Формирование двоичного кода осуществляется в направлении, обратном распространению сигнала при функционировании сети (т. е. от выходного слоя к входному):

Вид функции активации нейронов слоя	<i>tansig</i>	<i>logsig</i>	<i>purelin</i>
Шифр	00	01	10

Например, 101001 – входной слой имеет функцию активации *logsig*, скрытый – *purelin*, выходной – *purelin*.

Один из результатов экспериментов с трехслойными сетями представлен в таблице и на рис. 1, а. На рис. 2, б изображен общий

график зависимости отклонений прогноза от вариантов архитектуры и числа входных переменных.

Результаты, полученные в ходе опытов при одно-, двух- и трехслойных сетях.

Предельные относительные ошибки прогноза на тестовом множестве для однослойной и двухслойной сети (при отдельных вариантах комбинаций $\bar{x} + 3\sigma$ функций активации нейронов по слоям) находятся примерно на одном уровне и меньше, по отношению к ошибке прогноза трехслойной сети на 0,2–0,3 %.

Не все сочетания функций активации для нейронов входного, промежуточного и выходного слоев приемлемы с точки зрения точности выдаваемого результата. Для двух- и трехслойных сетей сочетание функций активации нейронов *purelin* для всех слоев дает минимальную величину отклонения.

Среди трехслойных сетей наилучшим образом обобщать результат на новые наблюдения способны сети, имеющие нейрон в выходном слое с линейной функцией активации, затем следуют сети с логарифмическими функциями активации, наихудшие результаты показывают сети с гиперболическим тангенсом в качестве функции активации выходного нейрона сети. Можно уверенно говорить о том, что сети, имеющие в выходном слое нейрон с линейной функцией активации, способны с достаточной степенью точности прогнозировать значение потребления электроэнергии.

Установлено, что число входных переменных слабо влияет на предельную относительную ошибку прогноза для первых пяти наблюдений. Для последующих наблюдений сеть с 16 входными переменными показывает лучшие результаты. Изменение числа предикторов меняет устойчивость сетей с точки зрения числа возможных вариантов их построения, дающих приемлемые результаты. Более устойчивыми оказались сети с 14 и 16 предикторами.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Барский А. Б. Нейронные сети: распознавание, управление, принятие решений. М.: Финансы и статистика, 2004. 176 с.
2. Круглов В. В., Борисов В. В. Искусственные нейронные сети. Теория и практика. М.: Горячая линия – Телеком, 2001. 382 с.
3. Рутковская Д., Пилинский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы / пер. с польск. И. Д. Рудинского. М.: Горячая линия – Телеком, 2006. 452 с.

4. Хайкин С. Нейронные сети. 2-е изд.; пер. с англ. Н. Н. Куссуль и А. Ю. Шелестова. М.: Изд. дом «Вильямс», 2006. 1104 с.
5. Карякин А. Л., Белов К. Д., Дегтярев Е. А. Прогнозирование потребления электроэнергии цехом агломерации на основе метода искусственных нейронных сетей // Изв. вузов. Горный журнал. 2008. № 8. С. 126–127.
6. Баврин И. И. Теория вероятностей и математическая статистика. М.: Высш. школа, 2005. 160 с.
7. Тейлор Дж. Введение в теорию ошибок: пер. с англ. М.: Мир, 1985. 272 с.

Поступила в редакцию 15 мая 2013 г.

Дегтярёв Евгений Андреевич – старший преподаватель кафедры электрификации горных предприятий. 620144, Екатеринбург, ул. Куйбышева, 30, Уральский государственный горный университет. E-mail: e.degtyarev@m.ursmu.ru

Карякин Александр Ливиевич – заведующий кафедрой электрификации горных предприятий, доктор технических наук, профессор. 620144, Екатеринбург, ул. Куйбышева, 30, Уральский государственный горный университет. E-mail: Aleksandr.Karyakin@m.ursmu.ru