



О.П. Солдатова, Я.А. Ледаков

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ РЕКУРРЕНТНОЙ СЕТЬЮ ЭЛЬМАНА

(ФГБОУ ВПО «Самарский государственный аэрокосмический университет им. академика С.П. Королева (национальный исследовательский университет)»)

Нейронные сети – это раздел искусственного интеллекта, в котором для обработки сигналов используются явления, аналогичные происходящим в нейронах живых существ [1].

Отдельную группу нейронных сетей составляют сети с обратной связью между различными слоями нейронов. Это так называемые рекуррентные сети. Их общая черта состоит в передаче сигналов с выходного либо скрытого слоя во входной слой.

Главная особенность таких сетей – динамическая зависимость на каждом этапе функционирования. Изменение состояния одного нейрона отражается на всей сети вследствие обратной связи типа «один ко многим». В сети возникает переходный процесс, который завершается формированием нового устойчивого состояния, отличающегося в общем случае от предыдущего.

Рекуррентная сеть Эльмана представляет собой развитие сетей персептронного типа за счет добавления в них обратных связей. Сеть Эльмана характеризуется частичной рекуррентностью в форме обратной связи между входным и скрытым слоем, реализуемой с помощью единичных элементов запаздывания z^{-1} . Обобщенная структура этой сети представлена на рисунке 1 [2].

Каждый скрытый нейрон имеет свой аналог в контекстном слое, образующем совместно с внешними входами сети входной слой. Выходной слой состоит из нейронов, однонаправленно связанных только с нейронами скрытого слоя.

Алгоритм обучения сети Эльмана можно представить в следующем виде:

1. Присвоить весам случайные начальные значения, имеющие, как правило, равномерное распределение в определенном интервале (например, между -1 и 1).

2. Для очередного момента t ($t = 0, 1, 2, \dots$) определить состояние всех нейронов сети (сигналы v_l и y_l). На этой основе можно сформировать входной вектор $x(t)$ для произвольного момента t .

3. Определить вектор погрешности обучения $e(t)$ для нейронов выходного слоя как разность между фактическим и ожидаемым значениями сигналов выходных нейронов.

4. Сформировать вектор градиента целевой функции относительно весов выходного и скрытого слоя.



5. Уточнить значения весов сети согласно правилам метода наискорейшего спуска.

6. После уточнения значений весов перейти к пункту 2 алгоритма для расчета в очередной момент времени t .

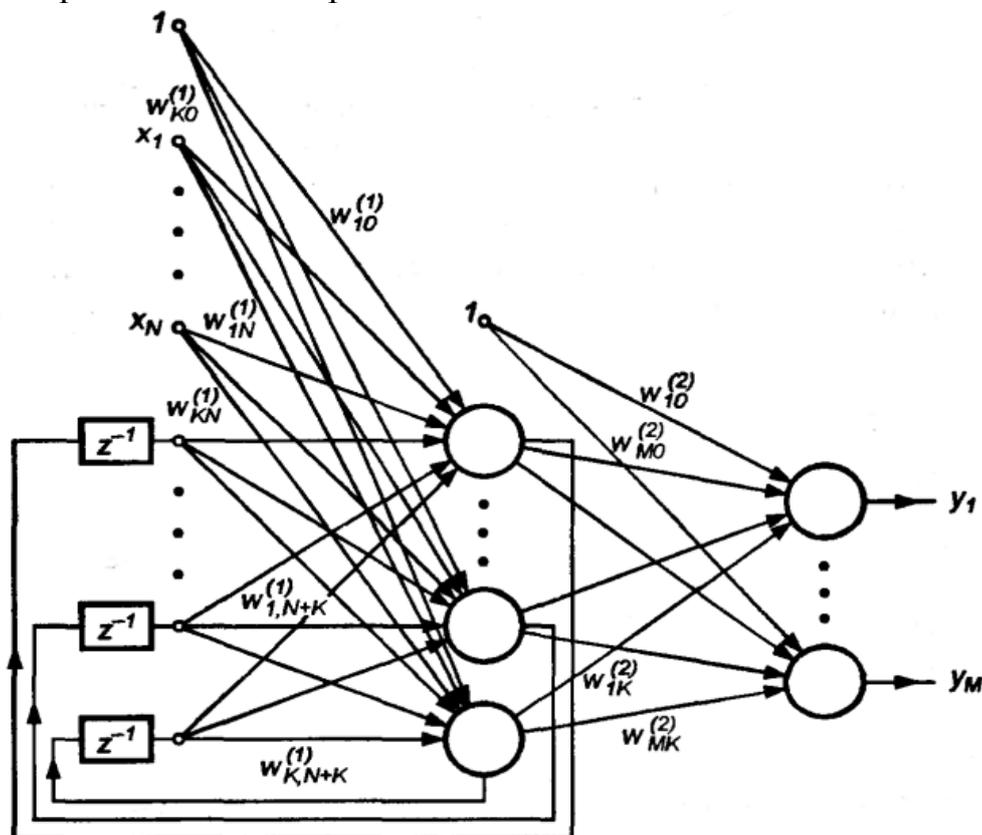


Рис. 1. Обобщенная структура сети Эльмана

Авторами разработан модуль для нейроимитатора «Нейрокомбайн», предназначенный для построения рекуррентной сети Эльмана. Для исследования эффективности решения задачи классификации была выбрана стандартная задача классификации вин. Набор данных представляет собой результаты химического анализа вин из винограда, выращенного в одном регионе, но принадлежащего к трем различным сортам. В ходе анализа выделено процентное содержание 13 составляющих вин, присутствующих в каждом из трех сортов. Задача – определить принадлежность вина с заданными характеристиками к одному из трех сортов. Общий объем данных – 178 образцов. Из них на 142 образцах проводилось обучение, а оставшиеся 36 использовались для тестирования сети.

В ходе экспериментальных исследований были определена зависимость среднеквадратического отклонения и средней приведенной погрешности от функции активации скрытого и выходного слоев при следующих параметрах сети: число нейронов во входном слое – 13, число нейронов в скрытом слое – 3, число нейронов в выходном слое – 1, число итераций обучения – 5000. Результаты исследования представлены в таблицах 1 и 2.



Таблица 1. Зависимость СКО погрешности от функции активации скрытого и выходного слоев

Скр. \ Вых.	Сигмоидальная	Логистическая	Ф. Гаусса	Синусоидальная
Сигмоидальная	0,103	0,126	0,142	0,137
Логистическая	0,097	0,121	0,075	0,114
Ф. Гаусса	0,113	0,107	0,102	0,086
Синусоидальная	0,105	0,125	0,087	0,116

Таблица 2. Зависимость средней приведенной погрешности от функции активации скрытого и выходного слоев

Скр. \ Вых.	Сигмоидальная	Логистическая	Ф. Гаусса	Синусоидальная
Сигмоидальная	11,78	14,05	16,38	15,37
Логистическая	10,78	14,44	7,85	13,90
Ф. Гаусса	13,35	13,74	11,09	07,54
Синусоидальная	12,51	15,22	9,30	13,84

Дальнейшее исследование проведено при следующих параметрах сети: число нейронов во входном слое – 13, число нейронов в выходном слое – 1, количество итераций обучения – 5000, функция активации для скрытого и выходного слоев - сигмоидальная. Результаты исследования представлены в таблице 3.

Таблица 3. Зависимость значений функции ошибки, СКО погрешности и средней приведенной погрешности от количества нейронов в скрытом слое

Количество нейронов в скрытом слое	Значение функции ошибки	СКО погрешности	Средняя приведенная погрешность
2	0,203	0,106	12,09%
3	0,229	0,113	12,75%
4	0,194	0,104	11,00%
5	0,259	0,120	12,95%
6	0,157	0,093	10,88%
7	0,136	0,087	08,14%
8	0,244	0,116	12,60%

На рисунке 2 представлены примеры тестирования на основе сетей с разными функциями активации.

Исходя из представленных данных, можно сделать вывод о том, что погрешность работы сети нелинейно зависит от количества нейронов в скрытом слое, и наблюдается тенденция уменьшения ошибки при увеличении количества нейронов скрытого слоя.

Так же стоит отметить хорошие результаты работы сети при использовании сочетания синусоидальной и гауссовой функций активации, а так же при использовании логистической функции активации скрытого слоя и гауссовой функции активации выходного слоя.

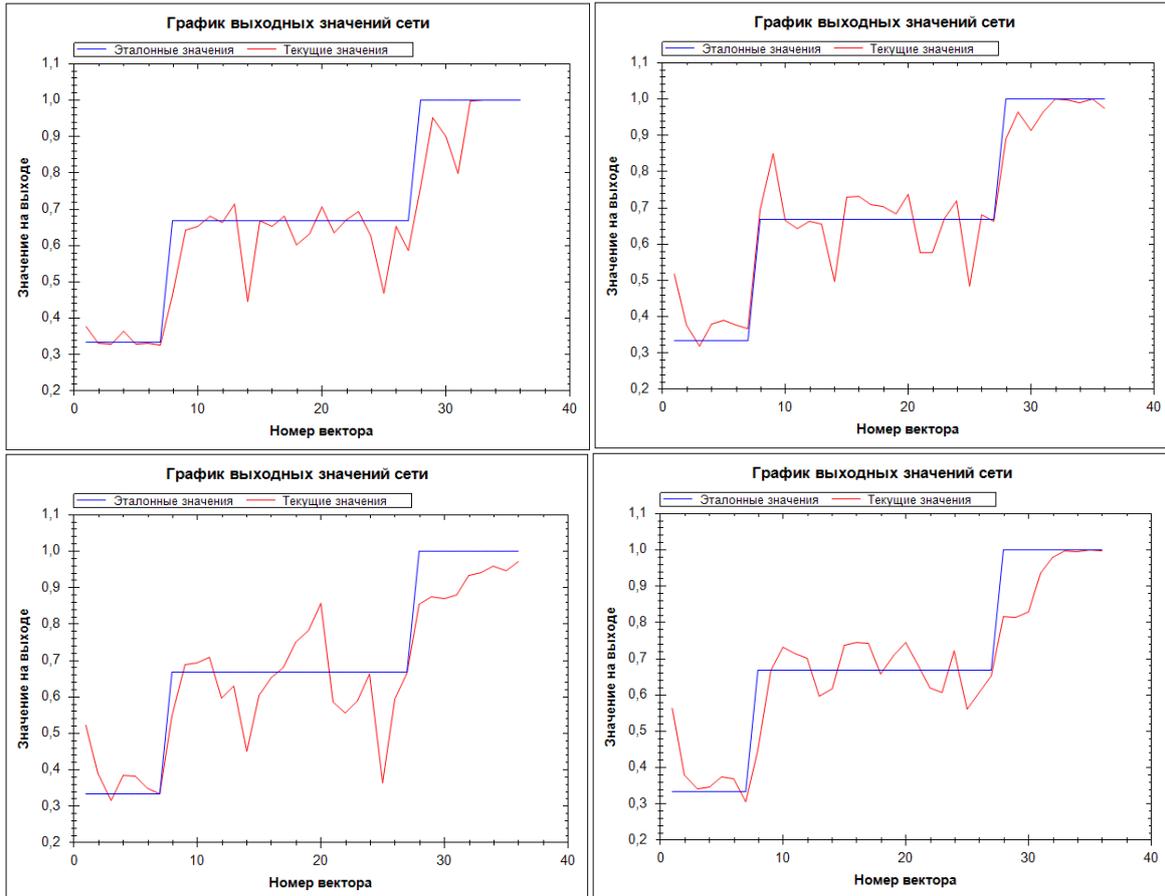


Рис. 2. Пример работы сетей с разными функциями активации

Литература

1. Хайкин С. Нейронные сети: полный курс. – М.: Вильямс, 2006.- 1104 с.
2. Осовский С. Нейронные сети для обработки информации / пер. с польского И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.