

ПРОЕКТИРОВАНИЕ НЕЙРОННОЙ СЕТИ В ТАБЛИЧНОМ ПРОЦЕССОРЕ MS EXCEL

Разработан алгоритм распознавания чисел с эмуляцией нейронной сети на основе использования стандартных функций табличного процессора MS Excel. Информация может быть полезной для специалиста и слушателя в плане изучения проблем нейронной сети.

Ключевые слова: нейронная сеть, алгоритм распознавания чисел.

Нейронные сети широко используются в распознавании образов, идентификации объектов и прогнозировании экономических процессов.

Под образом понимается структурированное описание изучаемого объекта, процесса или явления, представленное вектором признаков, выражаемых количественно. В результате идентификации объектов осуществляется их классификация, позволяющая выявлять закономерности и выполнять прогнозы [1].

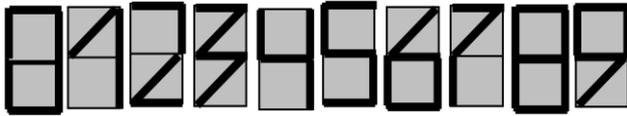
Нейронная сеть проектируется для решения узкоспециализированных задач в различных сферах [3]. Наиболее популярные и доступные приложения, позволяющие создавать нейронную сеть и управлять ее структурой – «Deductor Studio Academic», «NeuroEmulator», «NeuroView» и др., алгоритм обработки данных установлен разработчиками.

С целью наглядного отражения принципа работы нейронной сети спроектирован алгоритм распознавания знаков с использованием стандартных функций табличного процессора MS Excel. Решение компактно размещается на одной странице электронной книги (рисунок), обозримо без применения скроллинга [2] и позволяет ознакомиться с основными элементами и режимами работы нейронной сети в интерактивном режиме.

Системе предлагается распознать образы знаков десятичной системы (0, 1, 2, ..., 9), построенные с помощью горизонтальных, вертикальных и наклонных штрихов (коротких линий):

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	Распознавание (продолжение)																
2	Обучение						3 2 0 1				Σ	Порог					
3	Эталоны						Ошибка в разряде				1						
4		x_1	x_2	x_3	x_4		x_1	x_2	x_3	x_4							
5	0	2	4	0	0		1	2	0	1		4				1	
6	1	0	2	0	1		3	0	0	0		3				1	
7	2	2	1	1	0		1	1	1	1		4				1	
8	3	2	0	1	1		1	2	1	0		4				1	
9	4	1	3	0	0		2	1	0	1		4				1	
10	5	3	2	0	0		0	0	0	1		1	*			1	
11	6	2	2	0	1		1	0	0	0		1	*			1	
12	7	1	1	0	1		2	1	0	0		3				1	
13	8	3	4	0	0		0	2	0	1		3				1	
14	9	2	2	1	0		1	0	1	1		3				1	
15	Распознавание																
16	↓						Поток событий										
17																	
18	-3201						Элементарное событие										
19	3201						Фильтр										
20																	
21	↓						Рецепторы										
22																	
23																	
24																	
25	x_1	x_2	x_3	x_4													
26	3201	3201	3201	3201													
27	0,001	0,000	0,000	0,000													
28	0,000	0,010	0,000	0,000													
29	0,000	0,000	0,100	0,000													
30	0,000	0,000	0,000	1,000													
31	3,201	32,01	320,1	3201													
32	3	32	320	3201													
33																	
34																	
35																	
36	3 2 0 1																

Внешний вид рабочего листа задания



В частности, так выглядят табло индикаторных устройств, шаблоны почтовых индексов.



Например, образ знака «0» выполнен при помощи двух горизонтальных и четырех вертикальных штрихов.

Для идентификации образов полного набора (10 знаков) потребуется составить таблицу обучения нейронной системы по четырем признакам (видам штрихов), т. е. вектор i -го эталона X_i будет иметь следующие компоненты:

-  x_{i1} – количество вертикальных штрихов;
-  x_{i2} – количество горизонтальных штрихов;
-  x_{i3} – количество наклонных штрихов нижних;
-  x_{i4} – количество наклонных штрихов верхних.

Таблица эталонов располагается на листе MS Excel в виде матрицы (массив B5:E14).

Рассмотрим, как работает система.

Информация, поступающая с виртуального сканирующего устройства, представляет собой поток действительных чисел (X_i').

Рецептор (ячейка "B18") воспринимает (пропускает) числа по модулю < 10000 или сообщает о превышении допустимой разрядности.

В ячейку "B18" введено выражение:

=ABS(ЕСЛИ(ЦЕЛОЕ(B18)<=10000;ЦЕЛОЕ(B18);"Не более 4-х разрядов"))

Знак и дробная часть числа игнорируются.

«Чувствительность» рецепторов второго уровня определяет матрица "B27:E30".

Обработка выполняется последовательно в ячейках "B31:E31"; "B32:E32":

=МУМНОЖ(B26:E26;B27:E30);

=ОКРВНИЗ(C32;1), автозаполнение вправо.

На следующем уровне число (X_i') разносится по j разрядам в отдельные ячейки "B36:E36": из содержимого ячейки вычитается результат соседнего (старшего) разряда, умноженный на (-10):

=МУМНОЖ(B32:E32;G32:J35).

Для наглядности массив повторяется в ячейках "G6:J6", верхняя часть листа: содержимое ячеек соотносится с эталонами; суммарная ошибка (Функция $F_i = \sum(X_{ij}' - X_{ij})$) определяет степень адекватности регистрируемого события каждому из образов:

=ABS(B\$37-B6); автозаполнение вправо и вниз.

В качестве индикатора (*) используется логическое выражение:

=ЕСЛИ(L5<=N5;"*";" "), автозаполнение вниз.

На рабочем листе столбец "M5:M14" транспонирован в ячейки "N18:Q18" посредством "ТРАНСП".

Так как нейронные системы наиболее актуальны в решении нечетко формализованных задач, целесообразно задать предельную величину ошибки (порог), за которым событие будет отнесено к нераспознаваемому образу, т. е. не подлежащему классификации.

В нашем примере сигнал "3201", полученный с виртуального сканера, система относит либо к "5-ти", либо к "6-ти": в первом случае образ имеет лишний наклонный штрих, во втором – лишний вертикальный, суммарная ошибка равна "1".

Несмотря на то, что эталоны идентифицированы (соответствуют различным векторам X_i), некоторые ошибки выходят за рамки «осмысления» нейронной системой.



В частности, образ «безошибочно» распознается, как "9", в то время как вероятнее принять его как "6" (с неправильным положением наклонного штриха): алгоритм не учитывает положение горизонтальных и вертикальных штрихов, как и то, что они встречаются чаще наклонных.

«Аналитические» способности нашей системы, несомненно, возрастут, если функция F будет выражаться не в абсолютных величинах, а в частях; распознавание индексов улучшится при увеличении количества «нейронов» в слое. Отметим: при средней скорости восприятия информации, сопоставимой с человеческим зрением (30 кадров/с), система поддержит распознавание соответственно пяти-шестиразрядных индексов в секунду.

Следует учесть, что практически любой расчетный пример в MS Excel может быть эффективно дополнен средствами оптимизации.

Проектируемая система в принципе способна различать с требуемой точностью не только числа, но и текст.

В правой части рабочего листа представлен вариант конвертирования произвольного текста в последовательность чисел. Если в ячейку G25 помещено (импортируется, связывается) текстовое выражение "лето", разложение и перекодировка текста по ячейкам осуществляется командами =ЛЕВСИМВ(\$G25;1); =ПРАВСИМВ(G27;1); =КОДСИМВ(G28)-200 т. д.

Электронные таблицы интегрируются с базой данных и позволяют классифицировать любую информацию, в т. ч. экономическую – динамику цен, расчетные показатели, номенклатуру изделий, списки контрагентов (размерность массивов табличного процессора ограничивается в основном объемом оперативной памяти).

Настоящий пример наглядно отражает механизм работы нейронной сети (без использования операторов цикла, ветвления), опирается на общепринятые матричные процедуры и может быть рекомендован для освоения обучающимся по экономическим специальностям, маркетингу и менеджменту.

Библиографический список

1. Афонин, В. Л. Интеллектуальные робототехнические системы : курс лекций : учеб. пособие для студентов вузов, обучающихся по специальностям в обл. информ. технологий / В. Л. Афонин, В. А. Макушкин. – М. : Интернет-ун-т информ. технологий, 2005.
2. Денисов, Д. П. Рациональная организация расчетно-прикладных заданий в электронных таблицах MS Excel / Д. П. Денисов // Проблемы содержания и качества учебного процесса в высшей школе : сб. науч. ст. / под ред. А. Э. Еремеева, Г. В. Косякова. – Омск : Изд-во ОГИ, 2005. – С. 54–60.
3. Нейронные сети в MS Excel : методич. указания к практич. занят. и лаб. работам / сост. В. Х. Федотов ; Чуваш. ун-т. – Чебоксары, 2004. – 72 с.

© Денисов Д. П., Касимова О. К., 2010

Авторы статьи: **Дмитрий Павлович Денисов**, доцент кафедры ИМиЕНД, кандидат сельскохозяйственных наук (НОУ ВПО «ОмГА»);

Ольга Камалиеновна Касимова, специалист отдела дистанционного обучения (НОУ ВПО «ОмГА»).

Рецензент – Н. И. Ермолина, кандидат медицинских наук, доцент кафедры социальной работы и специальной психологии НОУ ВПО «ОмГА».