

куррентных сетей и НС на основе радиальных базисных функций (РБФ). Данные сети обладают определенными преимуществами по сравнению с сетями прямого распространения. Так, НС РБФ позволяют в процессе обучения адаптировать свою архитектуру и тем самым преодолеть недостатки многослойного перцептрона, связанные с проблемой выбора количества слоев и нейроэлементов в каждом слое. Рекуррентные сети обладают большими возможностями по способам организации обучающих множеств и адаптации прогнозных моделей. Для данных НС разработаны программные модели, на основе реальных временных рядов и математических функций исследованы прогнозирующие свойства, оптимизированы алгоритмы обучения по точности и скорости обучения. По результатам экспериментов делаются выводы по выбору структуры и типов НС, способам организации обучающих наборов на основе временных рядов, анализируются достоинства и недостатки исследуемых НС в задачах прогнозирования по сравнению с НС прямого распространения.

УДК 681.323:519.246

МОДИФИКАЦИЯ СТРУКТУРЫ НЕЙРОННОЙ СЕТИ ПРЯМОГО РАСПРОСТРАНЕНИЯ И АЛГОРИТМА ЕЕ ОБУЧЕНИЯ

Телятников Р.В., Спиридонов С.В.

Военная академия Республики Беларусь

В последние несколько лет наблюдается беспрецедентный рост активности в области теории искусственных нейронных сетей (НС), в идейном отношении близких к перцептрону Розенблатта. В докладе рассматривается один из подходов к модификации структуры перцептрона на основе использования РБФ-ячеек, а также приводится соответствующий алгоритм обучения (АО), отличающийся рядом достоинств по сравнению со стан-

3. Искусственный интеллект и нейронные сети

дартным АО (back-propagation error, BPE). Проверка работоспособности НС и эффективности АО осуществлялась при решении задачи распознавания образов.

Проблемы синтеза многослойных НС.

Существует множество спорных вопросов при проектировании нейронных сетей прямого распространения (НСПР) [1,2] - например:

- сколько необходимо использовать слоев в сети;
- сколько следует выбрать элементов в каждом слое;
- как сеть будет реагировать на данные, не включенные в обучающую выборку
- какой размер обучающей выборки (ОВ) необходим для достижения требуемой способности сети к обобщению;
- как определить момент окончания обучения сети.

Хотя многослойные НСПР широко применяются для классификации и аппроксимации функций [3], многие параметры приходится определять путем проб и ошибок. Существующие теоретические результаты дают лишь слабые ориентиры для выбора этих параметров в практических приложениях. В докладе приводится подробный анализ аспектов, усложняющих решение вышеперечисленных проблем.

Особенности обработки информации в НСПР.

В процессе распространения информации по сети после каждого слоя формируется все более сложное пространство фазовых координат (ПФК). Формирование выходного ПФК осуществляется в соответствии с целевой функцией (ЦФ) алгоритма обучения BPE:

$$E = \frac{1}{2} \sum_j^{N_{\text{вых}}} (Y_j - D_j)^2. \text{ Данный алгоритм, наряду с рядом положитель-}$$

ных качеств, обладает одним значительным недостатком. ЦФ алгоритма предполагает оптимизацию ПФК только на выходном слое сети и не контролирует эффективности формирования ПФК на внутренних слоях. Как следствие - велика вероятность попадания в локальные экстремумы ЦФ, а также сильная зависимость результата обучения от начального распределения весов связей. Под эффективностью формирования ПФК будем понимать степень достижения такой ЦФ, в соответствии с которой в гиперпространстве признаков максимизируется расстояние между точками (образами), принадлежащих разным классам и минимизируется расстояние между точками одного класса. Некоторые критерии оценки эффективности ПФК достаточно хорошо описаны в теории статистического распознавания образов [5]. Иными словами, необходимо решить задачу статистического анализа описания объектов из ОВ для фазовых координат выходного и обрабатываемых слоев сети, и по результатам данного анализа осуществлять коррекцию весовых коэффициентов связей.

Решение данной задачи и реализацию требуемой ЦФ предлагается осуществить путем введения в структуру НСПР нейронов на базе радиальной базисной функции (РБФ) [6]. Сети, содержащие РБФ-ячейки способны оценивать степень подобия подаваемого на вход описания хранящимся в памяти образам, осуществляя тем самым (в процессе обучения) кластеризацию образов по классам.

Мы будем использовать РБФ-ячейки в качестве нейронов выходного слоя. РБФ-ячейка, показанная на рис. 1 определяется как нейрон, с выходом:

3. Искусственный интеллект и нейронные сети

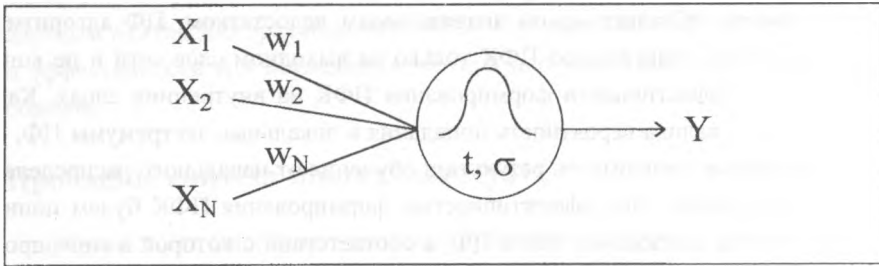


Рис. 1.

$$Y = N^{-1} \cdot \sum_{i=1}^N \exp\left(-|t_i - w_i \cdot x_i| / \sigma_i\right),$$

где t_i - математическое ожидание для i -ой входной координаты РБФ-ячейки, σ_i - дисперсия i -ой функции.

Алгоритм обучения модифицированной НСПР.

В процессе обучения первоначально определяются вектора математических ожиданий t ($t_i \in \mathbb{R}^N$) и дисперсий σ для элементов РБФ-слоя. Для определения дисперсии мы применим описанный в [6] алгоритм, согласно которому, σ_i пропорциональна расстоянию между образом t_i и его ближайшим соседом. Т.е. $\sigma_i = \tilde{\sigma} \cdot \min_{i \neq j} |t_i - t_j|$, где константа $\tilde{\sigma}$ выбрана *a priori*.

Далее корректируются веса входных связей обрабатывающих слоев (как и в алгоритме ВРЕ) в соответствии с выражением:

$$w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \eta \cdot \Delta w_{ij}.$$

Для обрабатывающего слоя, предшествующего РБФ-слою, Δw_{ij} рассчитывается по формуле:

$$\Delta w_{ij} = (D_j - Y_j) \cdot \sum_{k=1}^{N_{\text{РБФ}}} \Delta t_{jk} \cdot Y_i = \delta_j \cdot Y_i, \quad i = 1 \dots N_1, j = 1 \dots N_2, \quad (1)$$

связи 1-го обрабатывающего слоя корректируются по рекуррентной формуле:

$$\Delta w_{li} = \sum_{j=1}^{N_2} (\delta_j \cdot w_{ij}) \cdot F'_i \cdot Y_l = \delta_i \cdot Y_l, \quad l = 1 \dots N_{\text{Вх}}.$$

В выражении (1) Δt_{jk} представляет собой ошибку, характеризующую величину и знак отклонения j -ой фазовой координаты k -ой РБФ-ячейки от соответствующего математического ожидания t_{jk} данной ячейки. Если входной образ принадлежит r -му классу, то Δt_{jk} рассчитывается следующим образом:

$$\Delta t_{jk} = t_{jk} - w_{jk} \cdot Y_j, \quad j = 1 \dots N_2, k = 1 \dots N_{\text{РБФ}}, \quad \text{при } k = r,$$

$$\Delta t_{jk} = \text{sign}(t_{jk} - w_{jk} \cdot Y_j) \cdot (1 + |t_{jk} - w_{jk} \cdot Y_j|)^{-1}, \quad \text{при } k \neq r,$$

Результаты моделирования рассмотренной модифицированной НСПР показали, что применяемый для ее обучения АО имеет лучшую сходимость по сравнению с алгоритмом ВРЕ и, кроме того, позволяет выделить "способность обобщения сети" в отдельный регулируемый параметр.

Литература

1. Зарубежная радиоэлектроника. // М.: Радиоэлектроника. - № 5,6. - 1995.
2. Спиридонов С.В., Телятников Р.В. Комплексный алгоритм обучения искусственной нейронной сети прямого распространения. // Тез. докл.

IV Int. Conf. "Pattern Recognition and Information Processing". - Минск, 1997. - с.242.

3. J. Hertz, A. Krogh, R.G. Palmer. Introduction to the Theory of Neural Computation. // Addison-Wesley. Reading. Mass. 1991.
4. Веденов А.А., Ежов А.А., Левченко Е.Б. Архитектурные модели и функции нейронных ансамблей. // М.: ВИНТИ. - Итоги науки и техники. - Т.1. - "Физические и математические модели нейронных сетей". - 1990. - с.44-90.
5. Фукунага К. Введение в статистическую теорию распознавания образов. // Пер. с англ. - М.: Наука. Главная редакция Ф-М лит-ры. 1979.
6. Mark A. Neifeld, Demetri Psaltis. Optical implementations of radial basis classifiers. Applied Optics v. 32, n. 8, 1993.

УДК 621.865.8

ПОДСИСТЕМА ВИЗУАЛЬНОГО КОНСТРУИРОВАНИЯ МОДЕЛЕЙ НЕЙРОННЫХ СЕТЕЙ И ЭКСПЕРИМЕНТОВ

Костомаров В.В. Хвещук В.И.

Брестский политехнический институт

Подсистема визуального конструирования моделей нейронных сетей и экспериментов (ПВКМ) является составной частью системы автоматизации моделирования нейронных сетей. Идея визуального конструирования нейронных сетей (НС) родилась не случайно. Анализ имеющихся систем подобного рода показал, что большинство из них обладают рядом серьезных недостатков, таких как:

- возможность построения модели НС только заранее известных и изученных сетей, с жестко заданной топологией сети;
- отсутствие средств подключения самостоятельно разработанных процедур для реализации передаточных функций нейрона и связи;