ПОДХОДЫ К РЕШЕНИЮ ЗАДАЧИ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

А.С. Рудаков,
аспирант Московского государственного университета печати
E-mail: shmel_ras@mail.ru

В статье предлагается методика выбора рационального подхода к решению задачи прогнозирования временных рядов с помощью нейронных сетей между существующим подходом аппроксимации и предлагаемым подходом кластеризации, который только косвенно упоминается в области прогнозирования. Предлагаемая методика позволяет повысить эффективность проектирования и применения нейронных сетей для решения задач прогнозирования временных рядов путем определения на начальном этапе проектирования наиболее подходящей для решения поставленной задачи структуры сети и алгоритма обучения. Для выполнения этой задачи в основу методики положен анализ исходных обучающих данных, известных из условия поставленной задачи. В статье предлагается алгоритм для реализации методики.

Введение

Одной из наиболее распространенных областей применения нейронных сетей в настоящее время является область экономических и финансовых систем [1]. Это обусловлено одним из основных достоинств нейронных сетей — возможностью имитации процессов с большим количеством влияющих параметров, что является очень сложным или невозможным для традиционных методов в подобных условиях, например методов оптимизации. Однако, следствием широкого распространения нейронных сетей стала разработка большого количества разнообразных алгоритмов обучения нейронных сетей и их модификаций. Таким образом, актуальной является задача разработки метода выбора рационального подхода к решению задачи прогнозирования временных рядов с помощью нейронных сетей.

В большинстве случаев для решения задач прогнозирования с помощью нейронных сетей в настоящее время применяют подход аппроксимации функции. При использовании подхода аппроксимации функции, настраиваемые параметры сети в результате обучения принимают вид, соответствующий некоторой функции, представленной входными и выходными векторами обучающего множества [2]. Данный подход в основном применяется в задачах прогнозирования, в которых каждому конкретному входному вектору, представленному входными параметрами нейронной сети, соответствует конкретное значение прогнозируемого вектора, представленного выходными параметрами нейронной сети:

$$y_i = f(x_i), \quad (1)$$

где x_i — i-ый входной вектор, y_i — соответствующее значение прогнозируемого вектора;
$f(x_i)$ — прогнозирующая функция.

Подход аппроксимации функции имеет свои недостатки. В случае непериодической зависимости входных и выходных данных не исключена ситуация отрицательного результата обучения, из-за сложной формы аппроксимируемой функции.

Отрицательный результат обучения возможен также в условиях неполных данных, необходимых для успешной аппроксимации функции. Избежать подобных результатов можно, например, путем ввода в систему некоторой допустимой погрешности обучения.

Для решения задач прогнозирования временных рядов с помощью нейронных сетей предлагается также применять еще один подход, который только косвенно упоминается в области прогнозирования, — подход кластеризации. Применение подхода кластеризации предпочтительно в случаях, когда все множество входных векторов можно разбить на отдельные непересекающиеся подмножества — классы, или кластеры, и поставить в соответствие каждому подмножеству отдельный выходной вектор:

$$y^j = \delta(x_v),\ x_v \in X_j,\ j \in [1, k], \tag{2}$$

где x_v — v-ый входной вектор;

y^j — выходной вектор;

$\delta(x_v)$ — прогнозирующая функция;

X_j — подмножество обучающих входных векторов — класс, которое соответствует выходному вектору;

k — количество непересекающихся подмножеств входных векторов — классов и, соответственно, количество выходных векторов.

Таким образом, задача прогнозирования временных рядов с применением подхода кластеризации заключается в нахождении класса, к которому в большей степени подходит входной вектор, представленный значениями входных параметров нейронной сети. Подход кластеризации упрощает задачу прогнозирования, так как в нем происходит аппроксимация подмножеств обучающих входных векторов, а не отдельных обучающих входных векторов, что увеличивает скорость обучения, уменьшает время проектирования и повышает эффективность применения нейронной сети. Одним из недостатков подхода кластеризации является конечное число возможных прогнозируемых классов, которое зависит от количества выходных параметров проектируемой нейронной сети. Для увеличения числа возможных прогнозируемых классов можно увеличить размеры проектируемой нейронной сети.

Часто главная роль в постановке задачи прогнозирования отводится человеку: он определяет цель задачи, ее исходные данные и подход к решению. Положительный результат в решении поставленной задачи в большей степени определяется опытом и знаниями разработчика. Однако проектирование нейронной сети даже опытным специалистом не гарантирует рационального выбора подхода к решению задачи прогнозирования, из-за возможной человеческой ошибки, что подтверждает актуальность разработки метода выбора рационального подхода к решению задачи прогнозирования, позволяющего более обоснованно принять решение по выбору подхода. Выбор рационального подхода влияет на эффективность проектирования и применения нейронной сети для решения задачи прогнозирования временных рядов, так как на начальном этапе определяет наиболее подходящую для поставленной задачи структуру сети и алгоритм обучения.

Структуры и алгоритмы обучения нейронных сетей для реализации подходов к решению задач прогнозирования временных рядов

Каждый из подходов к решению задачи прогнозирования требует соответствующей структуры и алгоритма обучения нейронной сети. В большинстве случаев для решения задачи аппроксимации функции применяют алгоритм обратного распространения ошибки [2, 3, 4].

Алгоритм обратного распространения ошибки представляет собой итеративный градиентный алгоритм обучения, который используется с целью минимизации среднеквадратичного отклонения текущего выхода от желаемого выхода многослойной нейронной сети [2]. Обобщенная структура двухслойной нейронной сети для реализации алгоритма обратного распространения ошибки представлена на *рис. 1* [5], где w_{nm} — весовые коэффициенты связей между нейронами, OUT_m — выходные сигналы нейронов выходного слоя, n — количество нейронов в скрытом слое, m — количество нейронов в выходном слое. Входной слой выполняет только распределительные функции.

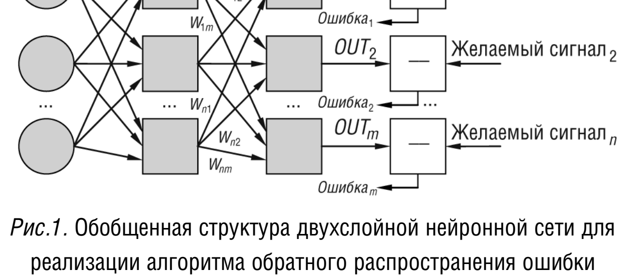

Рис.1. Обобщенная структура двухслойной нейронной сети для реализации алгоритма обратного распространения ошибки

Для подхода кластеризации предлагается применять структуру нейронной сети и алгоритм обучения Кохонена, так как данная структура и алгоритм применяются в основном для кластеризации входных образов [2, 3, 4]. Алгоритм обучения Кохонена также называют алгоритмом конкурентного обучения, а нейронные сети, использующие такой алгоритм — конкурентными. Конкурентные нейронные сети относятся к самоорганизующимся нейронным сетям. Самоорганизующиеся нейронные сети характеризуются обучением без учителя, в результате которого происходит адаптация сети к решаемой задаче. Структура конкурентной нейронной сети в общем случае представляет собой двухслойную нейронную сеть с прямыми связями. Первый слой выполняет распределительные функции, причем каждый нейрон первого слоя имеет связи со всеми нейронами второго слоя, который является выходным. Второй слой осуществляет конкуренцию между нейронами, в результате которой определяется нейрон-победитель. Победителем в конкуренции считается нейрон, который в результате подачи на вход сети определенного входного вектора имеет максимальную взвешенную активность. Для нейрона-победителя весовые коэффициенты усиливаются, а для остальных нейронов не изменяются или уменьшаются. По мере поступления входных векторов посредством обучения происходит разбиение -мерного входного пространства на различные области решений, каждой из которых соответствует отдельный нейрон выходного слоя [2]. Обобщенная структура конкурентной нейронной сети представлена на рисунке 2 [2], где w_{nm} — весовые коэффициенты связей между нейронами, OUT_m — выходные сигналы нейронов выходного слоя, n — количество нейронов в распределительном слое, m — количество нейронов в выходном слое.

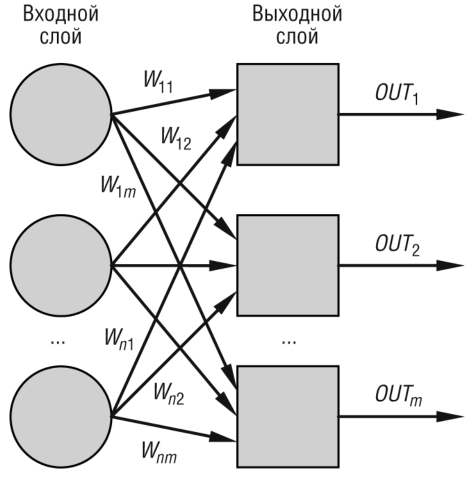

Рис.2. Обобщенная структура конкурентной нейронной сети

Методика выбора рационального подхода к решению задачи прогнозирования с помощью нейронных сетей

В основу предлагаемой методики выбора рационального подхода к решению задачи прогнозирования с помощью нейронных сетей между существующим и предлагаемым подходами положен анализ исходных обучающих данных, известных из условия поставленной задачи.

Пусть γ — вектор, размерность которого равна размерности входного вектора, содержащий информацию о минимально возможном расстоянии между подмножествами обучающих входных векторов — классами, и пусть выполнены следующие условия:

1. Подмножества X_j создают разбиение множества X, для которого верно:

$$X = \{X_1, X_2, ..., X_k\}, \tag{3}$$

$$X_j \cap X_l = 0,\ j \neq l,\ j \in [1,k]\ l \in [1,k], \tag{4}$$

$$\bigcup_{j=1}^{k} X_j = X, \tag{5}$$

где X — множество обучающих входных векторов;
X_j — j-ое подмножество обучающих входных векторов — класс;
k — количество подмножеств обучающих входных векторов — классов.

2. В каждом из подмножеств расстояние между наиболее близкими друг к другу соседними обучающими входными векторами — x_b^j и x_p^j — не превышает y:

$$\left|x_b^j - x_p^j\right| < \gamma, \tag{6}$$

где $j \in [1, k]$, k, — количество подмножеств обучающих входных векторов — классов;
$b \in [1, z_j]$, $p \in [1, z_j]$, z, — количество обучающих входных векторов в j-ом подмножестве обучающих входных векторов — классе;
x_b^j и x_p^j — наиболее близкие друг к другу соседние обучающие входные вектора j-го подмножества обучающих входных векторов — класса.

3. Расстояние между соседними подмножествами обучающих входных векторов — классов превышает или равно γ:

$$\left|x_d^j - x_r^h\right| \geq \gamma, \tag{7}$$

где $j \in [1, k]$, $h \in [1, k]$, $j \neq h$, k — количество подмножеств обучающих входных векторов — классов;

x_d^j и x_r^h — соответственно наиболее близкие друг к другу обучающие входные вектора соседних подмножеств обучающих входных векторов — классов;

$d \in [1, z_j]$, z_j — количество обучающих входных векторов в j-ом подмножестве обучающих входных векторов — классе;

$r \in [1, z_h]$, z_h — количество обучающих входных векторов в h-ом подмножестве обучающих входных векторов — классе.

4. Каждому обучающему входному вектору из j-го подмножества обучающих входных векторов — класса, где $j \in [1, k]$ и k — количество подмножеств обучающих входных векторов — классов, соответствует один выходной вектор — y^j:

$$\alpha_w^j = (x_w^j, y^j), \tag{8}$$

где $w \in [1, z_j]$, z_j — количество обучающих входных векторов в j-ом подмножестве обучающих входных векторов — классе;

$j \in [1, k]$, k , — количество подмножеств обучающих входных векторов — классов;

α_w^j — w-ая обучающая входная выборка j-го подмножества обучающих входных векторов — класса.

Тогда, при одновременном выполнении всех приведенных выше условий, все множество обучающих выборок представляется в виде отдельных подмножеств, каждому из которых соответствует определенный выходной вектор. Следовательно, в этом случае, для решения задачи прогнозирования с помощью нейронных сетей целесообразно применять подход кластеризации.

Таким образом, предлагается следующая методика выбора рационального подхода к решению задачи прогнозирования с помощью нейронных сетей, представимая в виде последовательности следующих этапов:

1. Проектировщиком нейронной сети, на основе своих знаний и опыта, задается вектор — вектор, размерность которого равна размерности входного вектора, содержащий информацию о минимально возможном расстоянии между подмножествами обучающих входных векторов — классами.

2. Осуществляется определение подмножеств обучающих входных векторов — классов X_j, создающих разбиение множества обучающих входных векторов X:

2.1. Для рассмотрения выбирается очередной обучающий входной вектор x_m, где $m \in [1, z]$, z — общее количество обучающих входных векторов.

2.2. Для рассмотрения выбирается очередной обучающий входной вектор x_q, где $q \in [1, z]$, z — общее количество обучающих входных векторов, $q \neq m$.

2.3. Проверяется выполнение условия на принадлежность пары обучающих входных векторов x_m и x_q к одному подмножеству обучающих входных векторов — классу:

$$\left| x_m - x_q \right| < \gamma . \tag{9}$$

Если условие (9) выполняется, идентификатор подмножества обучающего входного вектора x_m определен, $k_{x_m} \neq 0$, который вводится для каждого обучающего входного вектора и используется для хранения информации о принадлежности обучающего входного вектора к определенному подмножеству, и идентификатор подмножества обучающего входного вектора x_q не определен, $k_{x_q} = 0$, то $k_{x_q} = k_{x_m}$ и осуществляется переход к этапу 2.7.

Иначе, если условие (9) выполняется, $k_{x_m} = 0$ и $k_{x_q} \neq 0$, то $k_{x_m} = k_{x_q}$ и осуществляется переход к этапу 2.7.

Иначе, если условие (9) выполняется, $k_{x_m} = 0$ и $k_{x_q} = 0$, то увеличивается значение счетчика подмножеств k, содержащего информацию о количестве подмножеств обучающих входных векторов, $k_{x_m} = k$, $k_{x_q} = k$ и осуществляется переход к этапу 2.7.

Иначе, если условие (9) выполняется, $k_{x_m} \neq 0$, $k_{x_q} \neq 0$, то уменьшается значение счетчика подмножеств k и осуществляется переход к этапу 2.4.

2.4. Для рассмотрения выбирается очередной обучающий входной вектор x_t, где $t \in [1, z]$, z, — общее количество обучающих входных векторов, или выборок.

2.5. Если $k_{x_m} > k_{x_q}$ и $k_{x_t} = k_{x_m}$, то $k_{x_t} = k_{x_q}$.

Иначе, если $k_{x_m} > k_{x_q}$ и $k_{x_t} > k_{x_m}$, то $k_{x_t} = k_{x_t} - 1$.

Иначе, если $k_{x_m} < k_{x_q}$ и $k_{x_t} = k_{x_q}$, то $k_{x_t} = k_{x_m}$.

Иначе, если $k_{x_m} < k_{x_q}$ и $k_{x_t} > k_{x_q}$, то $k_{x_t} = k_{x_t} - 1$.

2.6. Если рассмотрены все обучающие входные вектора x_t, то осуществляется переход к этапу 2.7. Иначе осуществляется переход к этапу 2.4.

2.7. Если рассмотрены все обучающие входные вектора x_q, то осуществляется переход к этапу 2.8. Иначе осуществляется переход к этапу 2.2.

2.8. Если рассмотрены все обучающие входные вектора x_m, то осуществляется переход к этапу 3. Иначе осуществляется переход к этапу 2.1.

3. Осуществляется определение возможности применения обучающих входных выборок подмножеств обучающих входных векторов — классов в подходе кластеризации:

3.1. Для рассмотрения выбирается очередное подмножество обучающих входных векторов — класс X_j с номером j.

3.2. Для рассмотрения выбирается очередная обучающая входная выборка (x, y_i).

3.3. Если значение идентификатора подмножества выбранного обучающего входного вектора x_i равен значению номера выбранного подмножества j и данное равенство выполнилось впервые для выбранного подмножества обучающих входных векторов — класса, то $y^j = y_i$, где y^j — выходной вектор для j-го подмножества обучающих входных векторов — класса, и устанавливается флаг выбранной обучающей входной выборки (x_i, y_i), показывающий возможность ее применения в подходе кластеризации, $f_i = 1$, иначе, если значение идентификатора подмножества выбранного обучающего входного вектора x_i равен значению номера выбранного подмножества j и $y_i = y^j$, то $f_i = 1$.

3.4. Если рассмотрены все обучающие входные выборки (x_i, y_i), то осуществляется переход к этапу 3.5, иначе — к этапу 3.2.

3.5. Если рассмотрены все подмножества обучающих входных векторов X_j, то осуществляется переход к этапу 3.6, иначе — к этапу 3.1.

3.6. Если все обучающие входные выборки (x_i, y_i) имеют возможность применения в подходе кластеризации: $f_i = 1$, то для данного вектора y целесообразно применять подход кластеризации, иначе — подход аппроксимации.

Изменяя значения параметров вектора y, проектировщик нейронной сети имеет возможность получать различное количество подмножеств обучающих входных векторов. Чем меньше вектор y, тем большее количество подмножеств обучающих входных векторов можно получить. При выборе подхода кластеризации следует учитывать количество полученных подмножеств обучающих входных векторов, так как чем больше количество подмножеств, тем большим должно быть число выходных параметров проектируемой нейронной сети.

Алгоритм для реализации методики выбора рационального подхода к решению задачи прогнозирования с помощью нейронных сетей

Для реализации методики выбора рационального подхода к решению задачи прогнозирования с помощью нейронных сетей предлагается следующий алгоритм, который включает в себя следующие шаги:

1. Ввод проектировщиком нейронной сети вектора γ — вектора, размерность которого равна размерности входного вектора, содержащего информацию о минимально возможном расстоянии между подмножествами обучающих входных векторов — классами.

2. Инициализация переменных:

$$m = 1,\ q = 1,\ t = 1,\ i = 1,\ j = 1,\ k = 1,\ n = 0,$$

где $m \in [1, z]$, $q \in [1, z]$, $t \in [1, z]$, $i \in [1, z]$, z, — общее количество обучающих входных выборок;
$j \in [1, k]$, k, — счетчик, содержащий информацию о количестве подмножеств обучающих входных векторов — классов;
n — флаг, устанавливаемый при первом нахождении обучающей входной выборки определенного подмножества обучающих входных векторов.

где $i \in [1, z]$, k_{x_i}, — идентификатор подмножества обучающего входного вектора, который вводится для каждого обучающего входного вектора и используется для хранения информации о принадлежности обучающего входного вектора определенному подмножеству;

где $i \in [1, z]$, f_i, — флаг i-ой обучающей входной выборки, показывающий возможность ее применения в подходе кластеризации.

3. Выбор очередного обучающего входного вектора x_m.

4. Выбор очередного обучающего входного вектора x_q.

5. Проверка условия (9). Если условие (9) выполняется, $k_{x_m} \neq 0$ и $k_{x_q} = 0$, то $k_{x_q} \leftarrow k_{x_m}$ и осуществляется переход к шагу 9.

Иначе, если условие (9) выполняется, $k_{x_m} = 0$ и $k_{x_q} \neq 0$, то $k_{x_m} \leftarrow k_{x_q}$ и осуществляется переход к шагу 9.

Иначе, если условие (9) выполняется, $k_{x_m} = 0$ и $k_{x_q} = 0$, то $k \leftarrow k + 1$, $k_{x_m} \leftarrow k$, $k_{x_q} \leftarrow k$ и осуществляется переход к шагу 9.

Иначе, если условие (9) выполняется, $k_{x_m} \neq 0$, $k_{x_q} \neq 0$, то $k \leftarrow k-1$ и осуществляется переход к шагу 6.

6. Выбор очередного обучающего входного вектора x_i.

7. Если $k_{x_m} > k_{x_q}$ и $k_{x_t} = k_{x_m}$, то $k_{x_t} \leftarrow k_{x_q}$.

Иначе, если $k_{x_m} > k_{x_q}$ и $k_{x_t} > k_{x_m}$, то $k_{x_t} \leftarrow k_{x_t} - 1$.

Иначе, если $k_{x_m} < k_{x_q}$ и $k_{x_t} = k_{x_q}$, то $k_{x_t} \leftarrow k_{x_m}$.

Иначе, если $k_{x_m} < k_{x_q}$ и $k_{x_t} > k_{x_q}$, то $k_{x_t} \leftarrow k_{x_t} - 1$.

8. Если $t = z$, то осуществляется переход к шагу 9.

Иначе $t \leftarrow t+1$ и осуществляется переход к шагу 6.

9. Если $q = z$, то осуществляется переход к шагу 10.

Иначе $q \leftarrow q+1$ и осуществляется переход к шагу 4.

10. Если $m=z$, то осуществляется переход к этапу 11.

Иначе $m \leftarrow m + 1$ и осуществляется переход к шагу 3.

11. Выбор очередной обучающей выборки (x_i, y_i).

12. Если $k_{x_i} = j$ и $n = 0$, то $n \leftarrow 1$, $y' \leftarrow y_i$ и $f_i \leftarrow 1$.

Иначе $m \leftarrow m+1$, если $k_{x_i} = j$, $n = 1$, $y_i = y^j$, то $f_i \leftarrow 1$.

13. Если $i = z$, то осуществляется переход к шагу 14.

Иначе $i \leftarrow i + 1$ и осуществляется переход к шагу 11.

14. Если $i = k$, то осуществляется переход к шагу 15.

Иначе $i \leftarrow 1$, $n \leftarrow 0$, $j \leftarrow j + 1$ и осуществляется переход к шагу 11.

15. Если $\forall i, f_i = 1$, где $i \in [1, z]$, z — общее количество обучающих входных выборок, f_i — флаг обучающей входной выборки, показывающий возможность ее применения в подходе кластеризации, то для данного вектора γ целесообразно применять подход кластеризации, иначе — подход аппроксимации.

Алгоритм выбора рационального подхода к решению задачи прогнозирования с помощью нейронных сетей считается завершенным.

Выводы

Предлагаемая в статье методика позволяет осуществить выбор рационального подхода к решению задачи прогнозирования временных рядов с помощью нейронных сетей между существующим подходом аппроксимации и предлагаемым подходом кластеризации, который только косвенно упоминается в области прогнозирования. Выбор рационального подхода позволяет повысить эффективность проектирования и применения нейронных сетей для решения задач прогнозирования временных рядов путем определения на начальном этапе проектирования наиболее подходящей для решения поставленной задачи структуры сети и алгоритма обучения. В основу методики положен анализ исходных обучающих данных, известных из условия поставленной задачи, что позволяет более обоснованно принять решение по выбору подхода к решению задачи прогнозирования временных рядов с помощью нейронных сетей. ■

Литература

1. Данько, Т.П. «Системы искусственного интеллекта в разработке корпоративных маркетинговых стратегий», Журнал «Маркетинг в России и за рубежом», № 5 / Т.П.Данько, М.А.Ходимчук. М.: Финпресс. 2000.

2. Головко, В.А. «Нейронные сети: обучения, организация и применение» / В.А.Головко. М.: ИПРЖР. 2001.

3. Круглов, В.В. «Нечеткая логика и искусственные нейронные сети» / В.В.Круглов, М.И.Дли, Р.Ю.Голунов. М.: ФИЗМАТЛИТ. 2001.

4. Крисилов, В.А. «Представление исходных данных в задачах нейросетевого прогнозирования» / В.А.Крисилов, К.В.Чумичкин, А.В.Кондратюк. Одесса: ОНПУ. 2003.

5. Уоссермен, Ф. «Нейрокомпьютерная техника: теория и практика» / Ф.Уоссермен. Пер. с англ. Ю.А.Зуев, В.А.Точенов. 1992.