

Пронин С.В., канд. техн. наук. (ХНАДУ)

ЗАДАЧИ ПРОГНОЗИРОВАНИЯ С ПОМОЩЬЮ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Анотація: У статті розглядається підхід до вирішення завдань прогнозу за допомогою штучних нейронних мереж.

Ключові слова: штучна нейронна мережа, прогноз, алгоритм навчання

Аннотация: В статье рассматривается подход к решению задач прогноза при помощи искусственных нейронных сетей.

Ключевые слова: искусственная нейронная сеть, прогноз, алгоритм обучения

Abstract: This paper presents an approach to solving the problems of the forecast using artificial neural networks.

Keywords: artificial neural network, forecast, learning algorithm

Введение

Задачи прогнозирования решаются в самых разнообразных областях человеческой деятельности, таких как наука, экономика, производство и множество других сфер. Прогнозирование является важным элементом организации управления как чисто технических так и социально-экономических систем.

Развитие методов прогнозирования непосредственно связано с развитием информационных технологий, в частности, с ростом объемов хранимых данных и усложнением методов и алгоритмов прогнозирования. Вычислительные возможности современных информационных систем позволяют применять для решения задач прогноза аппарат искусственных нейронных сетей.

В данной статье рассматривается задача прогнозирования с помощью искусственных нейронных сетей. Показывается преимущество многослойных

нейронных сетей, которым для качественной работы нужно большее количество нейронов в слоях и большее количество элементов обучающей выборки.

Анализ литературы

Так как в нашем случае решается задача по сути прогнозирования временного ряда то для её решения более подходят математические и статистические методы. Остановимся кратко на данных методах.

Один из традиционных методов построения прогноза и обработки временных рядов – это построение регрессионной зависимости прогнозируемого параметра от времени. При использовании методов регрессии предполагается, что прогнозируемый случайный процесс является стационарным, т.е. в каждый промежуток времени случайная величина, характеризующая процесс, имеет вероятностное распределение и содержит неизменные по времени параметры.

Методы авторегрессии, когда прогнозируемое значение параметра зависит от определенной совокупности предыдущих значений параметров временного ряда, также строятся на предположении о стационарности процесса. Авторегрессия говорит о так называемом «эффекте памяти» для временного ряда, об инертности и стационарности прогнозируемого процесса и о сохранении сложившейся парадигмы на интервале прогнозирования. Методы авторегрессии применимы для коротких зависимостей, поэтому её использование для долговременного прогнозирования нецелесообразно [1].

Для того, чтобы убрать предпосылку о стационарности изучаемого процесса Энгл и Боллерслев разработали методы ARCH и GARCH [1-3]. Они предположили, что изучаемый процесс, не стационарен, но будущее значение волатильности процесса может быть спрогнозировано по ряду предыдущих значений волатильности, т.е. прогнозируется распределение ошибки прогнозирования. Дальнейшим развитием методов ARCH и GARCH выступают методы нейронных сетей, когда система прогнозирования в автоматическом режиме производит оценку параметров регрессии, уменьшая функцию ошибки. Но методы ARCH и GARCH перестают работать, когда изучаемый процесс резко изменяет свою тенденцию, т.е. происходит, так называемый, парадигмальный разрыв. В отличие от методов ARCH и GARCH нейронные

сети за счет реализованных в них механизмов самообучения лишены данных недостатков [4-10].

Выбор типа сети для решения задачи прогноза

При выборе архитектуры сети для прогнозирования обычно испытывается несколько конфигураций с разным количеством элементов. Исходя из того, что задача прогнозирования является частью случая задачи регрессии, то возможно ее решена следующими типами нейронных сетей многослойным персептроном (MLP), радиально-базисной сети (RBF), обобщенно-регрессионной сети (GRNN). Проанализируем основные свойства данных сетей. Сеть, построенная на радиальных базисных функциях (RBF), имеет промежуточная прослойка из радиальных элементов, каждый из которых воспроизводит гауссовой поверхности отклика. Поскольку эти функции нелинейные, то для моделирования любой произвольной функции отклика нет необходимости использовать более одной промежуточной слою: достаточно лишь взять оптимальное число радиальных элементов. RBF-сети имеют как ряд достоинств (компактность, быстрая обучаемость), так и недостатков. Например, с "групповым" представлением пространства модели связано неумение сетей RBF экстраполировать свои выводы за область известных данных: при удалении от учебной множества значение функции отклика быстро падает до нуля [4-9].

Многослойный персептрон (MLP) при обучении запоминает предлагаемые примеры, постепенно подстраивая свои внутренние параметры. Сети с общей регрессией запоминают примеры в буквальном смысле. Каждому примеру – отдельный нейрон в скрытом слое сети, а затем, при применении сеть сравнивает предлагаемый пример с примерами, которые она помнит. Самый простой вариант применения искусственных нейронных сетей в задачах прогнозирования - использование обычного персептрона с одним, двумя, или (в крайнем случае) тремя скрытыми слоями. При этом на входы нейронной сети обычно подается набор параметров, на основе которого (по мнению эксперта) можно успешно прогнозировать. Выходом обычно является прогноз сети на будущий момент времени.

Аппроксимация плотности вероятности с помощью ядерных функций является методологической основой для вероятностных (PNN) и обобщенно-регрессионных (GRNN) нейронных сетей. В этих сетях в точку расположения каждого учебного наблюдения содержится гауссова ядерная функция. Окончательная выходная оценка сети получается как взвешенное среднее выходов по всем учебным наблюдениям, где величины весов отражают расстояние от этих наблюдений до той точки, в которой производится оценивание. Таким образом, более близкие точки вносят больший вклад в оценку прогнозируемого процесса.

Первая промежуточная прослойка сети GRNN состоит из радиальных элементов, а вторая промежуточная прослойка содержит элементы, которые помогают оценить взвешенный выход нейронов первого слоя и состоит из двух нейронов. Обобщенно-регрессионная сеть обучается почти мгновенно, но может получиться большой и медленной. Как и сеть RBF, сеть GRNN не обладает способностью эффективно экстраполировать данные. Несмотря на то, что принцип обучения и применения таких сетей в корне отличается от обычных персептронов, внешне сеть используется таким же образом, как и обычный персептрон. Говоря другими словами, это совместимые архитектуры в том смысле, что в работающей системе прогнозирования можно заменить работающий персептрон на сеть с общей регрессией, и все будет работать. Не потребуется проводить никаких дополнительных манипуляций с данными.

Если персептрон при обучении запоминал предлагаемые примеры, постепенно подстраивая свои внутренние параметры, то сети с общей регрессией запоминают примеры в буквальном смысле. Каждому примеру – отдельный нейрон в скрытом слое сети, а затем, при применении сеть сравнивает предлагаемый пример с примерами, которые она помнит, и на основе этого сравнения выдаст ответ.

Отсюда следует основной недостаток такой архитектуры – когда объем информации велик, сеть станет слишком большой и будет медленно работать.

Ещё один недостаток таких сетей заключается в невозможности выстраивания линии прогноза. Поэтому такие сети можно использовать только

в случаях, когда наблюдаемый процесс устойчив, или, после декомпозиции данных [6].

Формирование архитектуры сети

Для решения поставленной задачи можно применить многослойную сеть с прямой передачей информации, образованная искусственными нейронами типа Адаптивный Линейный Элемент (ADALINE). Этот выбор определяется тем, что адалина является одним из простейших эффективных обучающихся нейронов. Структурно адалина состоит из адаптивного ассоциатора и активационной функции, в качестве которой предлагается использовать логистическую функцию активации.

Первый слой данной сети образован нейронами с обратными связями, на которые подается задержанные значения входного сигнала, а также задержанные сигналы предсказанного значения.

При задании входного вектора $X(k)$ ретроспективная выборка будет приобретать вид:

$$X=X(k), X=X(k-1), X=X(k-2), \dots, X=X(k-n), \quad (1)$$

где $X(k)$ – входной вектор, характеризующий значение показателя в момент времени k ;

$X(k-n)$ – входной вектор, характеризующий значение показателя в момент времени $k-n$.

Таким образом ретроспективная выборка составит:

$$X(k) \rightarrow D(k), \quad (2)$$

где $D(k)$ – ожидаемое значение.

Вопрос количества промежуточных (скрытых) слоев и элементов в них решается экспериментальным путем, при этом количество промежуточных слоев и количество элементов в них выбирались таким образом, чтобы для различных наборов реализаций достигался минимум критериев:

- максимальной относительной ошибки прогноза;
- времени обучения (переобучения) нейронной сети

Оценка прогноза осуществляется по методу наименьших квадратов, который представляет собой метод оценки параметров модели на основании экспериментальных данных, содержащих случайные ошибки. Так необходимо минимизировать разницу между экспериментальными данными и теоретическими (вычисленными при помощи предложенной модели). Мерой разницы в методе наименьших квадратов служит сумма квадратов отклонений действительных (экспериментальных) значений от теоретических.

Выбор алгоритма обучения нейронной сети

На следующем этапе необходимо подобрать алгоритм обучения нейронной сети с целью получения более точного прогноза.

В основном все алгоритмы обучения можно разбить на две категории [4, 8-9]:

- градиентные алгоритмы;
- стохастические алгоритмы.

Первую группу составляют алгоритмы основанные на вычислении производной функции ошибки и корректировке весов в соответствии со значением найденной производной. Основу всех этих алгоритмов составляет хорошо известный алгоритм обратного распространения ошибки [4].

Стохастические методы обучения выполняют псевдослучайные изменения величин весов, сохраняя те изменения, которые ведут к улучшениям характеристик сети [6, 8-9]. Основным недостатком стохастических алгоритмов является достаточно долгое время обучения.

В статье, для выбранной архитектуры нейронной сети, в качестве алгоритма обучения нейронной сети был выбран алгоритм обучения Уидроу-Хоффа [10], обладающий высоким быстродействием.

В данном алгоритме входной сигнал это сумма $x_n(k)$ представляющий из себя входное значение прогнозируемого процесса $x_n(k)$ умноженную на вес $w_n^{[n]}$, который устанавливается в зависимости от характеристик наблюдаемого процесса:

$$\hat{x}(k) = w_1^{[1]}(k)x_1(k) + w_2^{[1]}(k)x_2(k) + \dots + w_n^{[1]}(k)x_n(k), \quad (3)$$

где $\hat{x}(k)$ – значение на выходе первого слоя нейронной сети;

$w_i^{[1]}(k)$ – значение синаптических весов первого слоя нейронной сети;

$x_i^{[1]}(k)$ – значение на входе нейронной сети.

Выход нейрона сравнивается с внешним воздействием $d(k)$, которое в данном случае является корректирующим коэффициентом в результате чего рассчитывается ошибка и корректировка весовых коэффициентов:

$$e^{[1]}(k) = d(k) - \hat{x}(k), \quad (4)$$

где $e^{[1]}(k)$ – значение погрешности обучения нейронной сети;

$d(k)$ – корректирующее значение.

$$w_i^{[1]}(k+1) = \omega_i^{[1]}(k) + \eta(k)e^{[1]}(k)x_i(k), i = 1 \dots n, \quad (5)$$

где $w_i^{[1]}(k+1)$ – скорректированное значение синаптических весов;

μ – коэффициент скорости обучения.

Далее выходной сигнал подается на цепочку из элементов чистого запаздывания z^{-1} в результате чего на выходе образуется последовательность:

$$\hat{y}(k) = w_1^{[2]}(k)\hat{x}(k-1) + w_2^{[2]}(k)x(k-2) + \dots + w_m^{[2]}(k)x(k-m) + w_{m+1}^{[2]}\hat{y}(k-1) + w_{m+2}^{[2]}\hat{y}(k-2) + \dots + w_{m+1}^{[2]}\hat{y}(k-1), \quad (6)$$

где $\hat{y}(k)$ – выходной сигнал сети;

$\hat{x}(k-1)$ – значение на выходе скрытого слоя нейронной сети;

$w_i^{[1]}(k+1)$ – скорректированное значение синаптических весов для второго слоя нейронной сети.

Выход нейрона сравнивается с сигналом $\hat{x}(k)$ после чего происходит подстройка синаптических весов выходного слоя по формуле:

$$w_j^{[2]}(k+1) = \omega_j^{[2]}(k) + \eta(k)e^{[2]}(k)\tilde{x}_j(k). \quad (7)$$

После чего выходным нейроном формируется прогноз.

Вывод

В статье был предложен подход к прогнозированию с помощью аппарата искусственных нейронных сетей. Определена архитектура сети и подобран алгоритм обучения.

Литература

1. Митропольский А. К. Техника статистических вычислений / А. К. Митропольский – М., Наука, 1971. – 576 с.
2. Учебник. Методы прогнозирования: Регрессионные методы прогнозирования [Электронный ресурс] / Аналитические технологии для прогнозирования и анализа данных. НейроПроект: 1999-2005. Режим доступа: URL: http://www.neuroproject.ru/forecasting_tutorial.php
3. Учебник. Методы прогнозирования: Регрессионные методы прогнозирования [Электронный ресурс] / Аналитические технологии для прогнозирования и анализа данных. НейроПроект: 1999-2005. Режим доступа: URL:
4. http://www.neuroproject.ru/forecasting_tutorial.php
5. Руденко О. Г., Искусственные нейронные сети / О. Г. Руденко, Е. В. Бодянский – ООО «Компания СМИТ», 2005. – 408 с.
6. Wasserman P. D. Neural Computing: Theory and Practice / Van Nostrand Reinhold, New York, NY 1989. – 189 с.
7. Нейросетевой анализ и прогнозирование: [Электронный ресурс]. Режим доступа: URL: <http://www.netneuro.ru>
8. Aarts E. H. L.. Simulated annealing and Boltzmann machines. Chichester : Wiley, 1989. – pp 443 – 447.
9. Медведев В. С. Нейронные сети. MATLAB 6 / В. С. Медведев, В. Г. Потемкин – М. : ООО «ВИТРЭМ», 2002. – 245 с.
10. Калацкая Л. В. Нейронные сети Statistica neural Networks / Л. В. Калацкая,
11. В. А. Новиков, В. С. Садков – М.: Телеком, 2000. Организация и обучение искусственных нейронных сетей / Мн.: БГУ – 2003. – 267 с.
Пронин С. В. Применение искусственных нейронных систем для моделирования транспортных систем / С. В. Пронин// Автомоб. трансп.: сб. науч. тр. – Х., 2006. – Вып.18. – С. 123 – 125.