

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ С АДАПТИВНОЙ ЦЕЛЕВОЙ ФУНКЦИЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

Брезинский Д.И., Серая О.В.
Национальный технический университет
«Харьковский политехнический институт»

Способность к моделированию нелинейных процессов, работе с зашумленными данными и адаптивность дают возможности применять нейронные сети для решения широкого класса задач, в том числе и прогнозирования временных рядов [1]. Нейронные сети в таких задачах - серьёзная альтернатива регрессионному анализу, традиционно используемому в этих случаях. Дело в том, что в ситуациях, когда на результирующий фактор влияет большое количество факторов и, возможно, их взаимодействия, размерность задачи становится неприемлемо высокой. Качество построенной нейросети зависит от качества и объёма используемых статистических данных, а также от вида активационных функций, применяемых для её построения [2].

Предложенная методика использования адаптивной активационной функции даёт дополнительные возможности повышения эффективности нейронных сетей.

В реальных нейронных сетях наиболее часто используются следующие активационные функции: пороговая униполярная, пороговая биполярная, логистическая, шаговая (линейная с насыщением), синусоидальная с насыщением и экспоненциальная. Используемые на практике некоторые другие активационные функции могут быть выражены через уже перечисленные.

В данной методике предлагается универсальная активационная функция, зависящая от некоторых параметров, при изменении которых она может принимать вид любой из перечисленных ранее. Данная функция имеет вид:

$$j_0(s) = b_0 + b_1 th(c_0 + c_1 s + c_2 s^2) ,$$

где b_0, b_1, c_0, c_1, c_2 - изменяемые параметры.

Адаптирующаяся активационная функция принимает вид логистической функции при таких значениях изменяемых параметров: (0,5;0,5;0; $\alpha/2$;0); экспоненциальной при (28;29;-1,7;-0,14;0,02); синусоидальной при (0,02;1,17;-0,01;1,4;-0,04); шаговой при (0,05;1,05;-0,02;1,23;-0,22); пороговой униполярной при (0,5;0,5;13;111;-21) и пороговой биполярной при (0;1;-12,5;146;-25).

Таким образом, проведенное исследование показало, что универсальное описание активационной функции обеспечивает хорошую аппроксимацию всех основных активационных функций. Это обстоятельство указывает на то, что при включении процедуры выбора параметров активационной функции в общую процедуру обучения нейронной сети можно не только автоматизировать выбор активационной функции, но и получить новый вид активационной функции, использование которого позволит лучше описать моделируемый процесс и получить более точные результаты при прогнозировании.

С помощью универсальной функции обучение сети может быть организовано с помощью следующей итерационной процедуры. На каждой итерации

решаются задачи двух типов – внутренние и внешние. Внешняя задача реализует шаг улучшения параметров адаптивной активационной функции путем реализации процедуры Нелдера-Мида. При этом в качестве целевой функции процедуры Нелдера-Мида используется значение ошибки выхода, получаемой при сравнении выхода сети с истинным, задаваемым при обучении. Совокупность внутренних задач обеспечивает настройку сети для каждого из наборов параметров активационной функции.

Рассмотрим процедуру обучения более подробно. В качестве начального симплекса из шести вершин, применяемого в методе Нелдера-Мида, используется матрица, состоящая из шести строк. Каждая строка – это значения параметров адаптирующейся функции, при которых она принимает вид одной из стандартных активационных функций. Далее для каждой активационной функции находится значение целевой функции задачи. Расчет этих значений происходит следующим образом. Выбирается конкретный вариант описания активационной функции, задающий соответствующую нейронную сеть. Далее заданный набор исходных данных используется для обучения сети путем выбора системы весовых коэффициентов, обеспечивающих наилучшую в этих условиях нейронную сеть. Процедура повторяется для всех вариантов активационных функций. При этом получен ряд значений, каждое из которых представляет собой ошибку выхода для соответствующей активационной функции. Теперь в соответствии со стандартной процедурой Нелдера-Мида осуществляется упорядочение полученных значений в порядке возрастания, выделение наилучшей точки с максимальным значением ошибки. Дальнейшие действия выполняются в полном соответствии со стандартным алгоритмом Нелдера-Мида. В результате этих действий происходит формирование нового симплекса. У этого симплекса будет одна новая вершина при отражении, растяжении и сжатии, или пять новых вершин, если произошла редукция. Первая итерация на этом закончена. На второй и всех последующих итерациях процедура повторяется. Вычисления продолжаются до выполнения какого-либо естественного критерия останова. В результате работы алгоритма находятся такие параметры адаптирующейся активационной функции, при которых ошибка выхода нейронной сети минимальна, что в свою очередь означает наиболее адекватное описание данных между входным набором переменных и результирующей переменной.

Особенностью данной методологии является то, что она позволяет настраивать конкретный вид активационной функции для всей сети, для каждого слоя или даже для каждого отдельного нейрона, но, естественно, это потребует значительно более сложных вычислений.

Предложенная процедура построения нейронной сети, использующая адаптивную активационную функцию с настраиваемыми параметрами, обеспечивает повышение эффективности использования сетей в задачах с ограниченным набором данных, что позволяет избежать переобучения сети, то есть запоминания обучающего набора данных. Параметрическое задание активационной функции позволяет осуществить её настройку, включив её в общую процедуру обучения.

Таким образом, получена технология формирования нейронной сети, реализующая глубокую её адаптацию по параметрам активационной функции и численным значениям системы весовых коэффициентов.

Литература

3. Руденко О.Г., Бодянский Е.В. Основы искусственных нейронных сетей. – Х.: Телетех, 2002. – 317с.
4. Бриллинджер Д. Временные ряды. Обработка данных и теория: Пер. с англ. – М.: МИР, 1980. – 536с.