УДК 621.31:51

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ МОДЕЛИРОВАНИЯ НЕЛИНЕЙНЫХ ЗАВИСИМОСТЕЙ В ЗАДАЧЕ ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ

Валь П.В., Соловьев Е.С.
Научный руководитель – профессор Попов Ю.П.

Сибирский федеральный университет

С созданием оптового рынка электроэнергии и мощности (ОРЭМ), формированием рынка на сутки вперед (РСВ) и балансирующего рынка (БР) необходимым условием нормального функционирования предприятий-участников рынка становится наличие методики почасового краткосрочного прогнозирования электропотребления.

Электропотребление промышленного предприятия зависит от множества технологических, производственных, климатических, горно-геологических, и других факторов. Большинство классических методов прогнозирования основывается на использовании линейной зависимости между прогнозируемой (зависимой) переменной (электропотребление) и объясняющими (независимыми) переменными (например, температура окружающего воздуха, план выпуска продукции и т.п.). Так, модель классической линейной множественной регрессии имеет вид

$$Y = a_0 + a_1 \cdot X_1 + a_2 \cdot X_2 + \dots + a_n \cdot X_n + e,$$

где: Y – оценка зависимой переменной; X_1, X_2, $\dots X_n$ – значения независимых переменных (факторов); a_0, a_1, a_2, $\dots a_n$ – коэффициенты линейной регрессии; e – случайная ошибка.

Построение моделей линейной множественной регрессии основывается на совместном использовании корреляционного и регрессионного анализа. Задача корреляционного анализа сводится к измерению тесноты связи между варьирующими признаками, определению неизвестных причинных связей и оценке факторов, оказывающих наибольшее влияние на результативный признак. Задачи регрессионного анализа лежат в сфере установления формы зависимости, определения функции регрессии, использования уравнения для оценки неизвестных значений зависимой переменной.

Широкое распространение методов корреляционного и регрессионного анализа в сфере анализа статистических данных главным образом обусловлено относительной простотой подсчета коэффициентов корреляции и регрессии. Коэффициент корреляции Пирсона, определяющий характер взаимного стохастического влияния изменения двух случайных величин, определяется по формуле

$$r = \frac{\sum\left((X_i - \bar{X}) \cdot (Y_i - \bar{Y})\right)}{\sqrt{\sum(X_i - \bar{X})^2} \cdot \sqrt{\sum(Y_i - \bar{Y})^2}},$$

где: $\bar{X}$, $\bar{Y}$ – средние значения величин X, Y.

На практике нередко возникают ситуации нелинейной зависимости между переменными, однако классический корреляционный анализ отражает лишь линейную зависимость величин. В качестве примера с помощью корреляционного анализа установим степень связи между величинам $A = sin(x)$ и $B = cos(x)$ и определим на основе регрессионного анализа форму регрессии $B = f(A)$. На рис. 1 представлена диаграмма рассеяния величин A и B (построения производились по 21 точке для значений x интервала от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$) с изображением линии регрессии $B = f(A)$, а также приведены коэффициент корреляции и уравнение регрессии (расчеты и построения произведены в программном пакете Statistica).

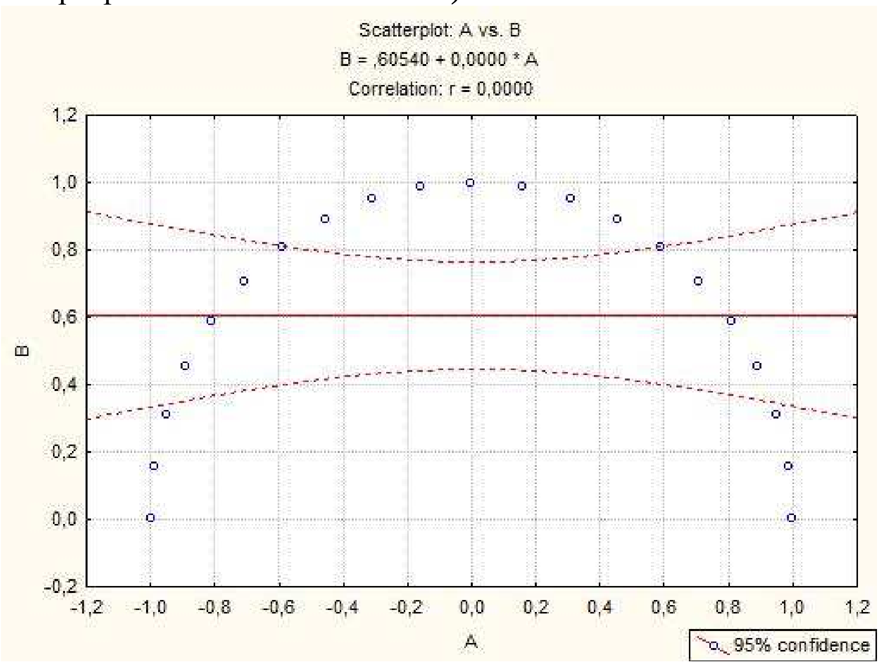

Рис. 1. Результаты корреляционного и регрессионного анализа рассматриваемого примера

Как видно из рис. 1 коэффициент корреляции равен нулю, а линия регрессии $B = f(A)$ параллельна оси абсцисс A, т.е. линейная зависимость между величинами отсутствует. Между тем, величины A и B очевидно связаны функционально по закону

$$sin^2(x) + cos^2(x) = 1.$$

Традиционными методами моделирования нелинейных зависимостей между переменными является подгонка кривых (логарифмических, гиперболических, квадратичных, кубических, степенных, показательных, логистических, экспоненциальных и др.), использование фиктивных переменных (нелинейных либо по переменным, либо по параметрам), используемых для приведения к линейной регрессии, и другие. Однако в настоящее время более перспективным методом моделирования нелинейных зависимостей в задачах прогнозирования является использование искусственных нейронных сетей.

Теория искусственных нейронных сетей (ИНС) возникла из исследований в области искусственного интеллекта. ИНС основаны на биологической модели нервной системы, составлены из множества простых элементов, действующих параллельно. Как и в природе, функция нейронной сети в значительной степени определяется связями между элементами. Нейронную сеть можно обучать для выполнения конкретной функции, регулируя значения коэффициентов (весов) связи. Обычно ИНС настраиваются или обучаются так, чтобы конкретные входы преобразовывались в заданный целевой выход. Сеть настраивается (обучается), основываясь на сравнении сигналов выхода и цели до тех пор, пока выход сети не будет соответствовать цели (обучение с учителем). Чтобы обучить сеть при таком управляемом обучении, как правило, используется много пар значений сигналов вход-цель.

Наибольшее применение на практике получили нейронные сети прямой передачи сигнала типа многослойный персептрон с одним или несколькими скрытыми слоями, обучение которых производится методом обратного распространения ошибки.

На рис. 2 представлен многослойный персептрон типа 3-4-1 (с тремя нейронами входного слоя, четырьмя нейронами скрытого и одним нейроном выходного).

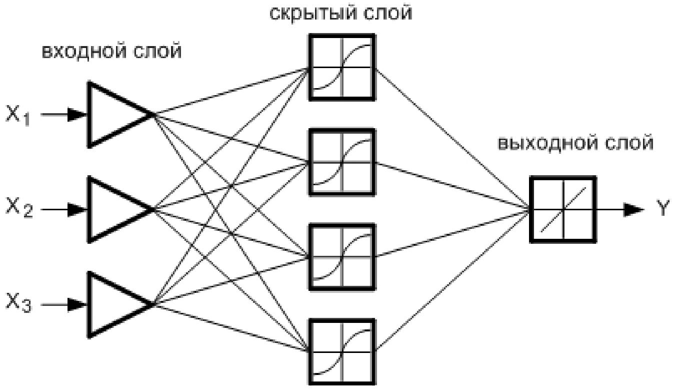

Рис. 2. Многослойный персептрон

Математически нейрон представляет собой взвешенный сумматор, единственный выход которого определяется через его входы и значения синаптических весов входов. Например, выход первого нейрона скрытого слоя ИНС (рис. 3) определяется выражением

$$Y_{hl1} = F_{hl1}(X_1 \cdot w_{11} + X_2 \cdot w_{21} + X_3 \cdot w_{31} + b_{hl1}),$$

где: X_1, X_2, X_3 – значения входных переменных; w_{11}, w_{21}, w_{31} – значения весов между нейроном и нейронами входного слоя; b_{hl1} – значение смещения нейрона; F_{hl1} – функция активации нейрона.

В качестве функций активации нейронов скрытого слоя обычно используются сигмоидные функции: логистическая (несимметричная) либо гиперболический тангенс (симметричная). Функция гиперболического тангенса представлена на рис. 3 и определяется выражением

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

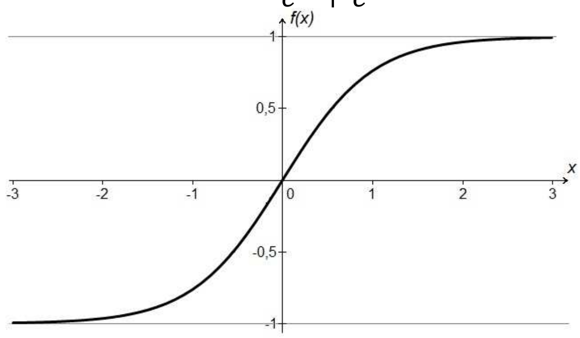

Рис. 3. Функция гиперболического тангенса

В качестве функции активации нейронов выходного слоя обычно используется линейная функция. Именно благодаря использованию нелинейных функций активации нейронов ИНС способны воспроизводить нелинейные зависимости.

Покажем, как с помощью многослойного персептрона (см. рис. 2) решить задачу воспроизведения нелинейной зависимости $B = f(A)$, используя программный пакет MATLAB (Neural Network Toolbox).

Для решения данной задачи, был использован описанный ранее многослойный персептрон со структурой 1-2-1. Обучающая выборка, включающая входной вектор (P) и целевой (T), состояла из 11 примеров:

$P = [-1\ -0.95\ -0.81\ -0.59\ -0.31\ 0\ 0.31\ 0.59\ 0.81\ 0.95\ 1]$;

$T = [0\ 0.31\ 0.59\ 0.81\ 0.95\ 1\ 0.95\ 0.81\ 0.59\ 0.31\ 0]$.

Обучение нейронной сети осуществлялось с помощью итеративного градиентного алгоритма Левенберга-Марквардта (Levenberg-Marquardt) в течение 50 итераций (эпох). На рис. 4 представлен график обучения, из которого видно, что нейронная сеть практически полностью обучилась к пятой эпохе, достигнув значения среднеквадратичной ошибки 0,0018.

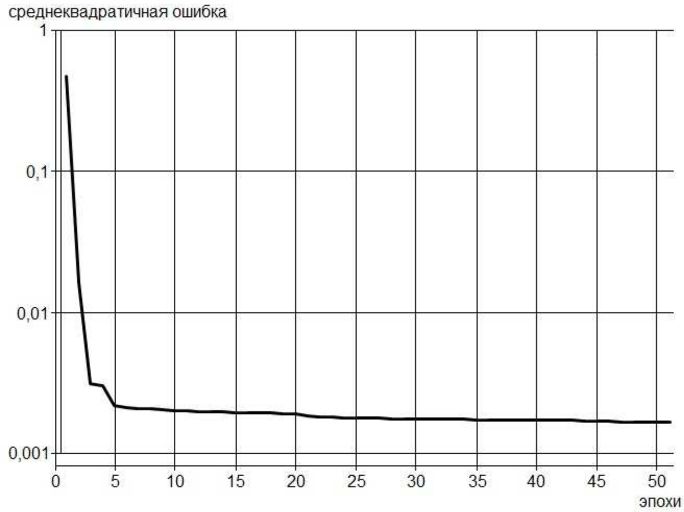

Рис. 4. График обучения многослойного персептрона

После обучения ИНС можно оценить качество аппроксимации нелинейной функции $B = f(A)$ посредством подачи на вход сети соответствующих значений величины A. На рис. 5. представлены результаты моделирования данной функции.

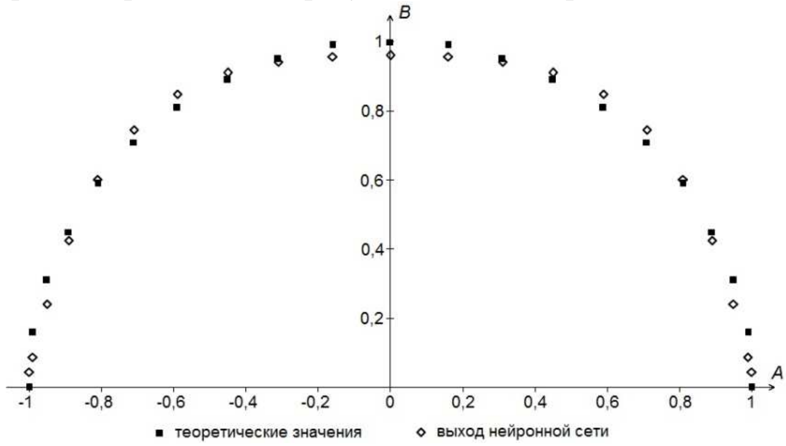

Рис. 5. Результаты моделирования $B = f(A)$ с помощью ИНС

Результаты моделирования нелинейной зависимости $B = f(A)$ свидетельствуют о том, что достаточно простая искусственная нейронная сеть довольно точно справилась с задачей, решение которой оказалось затруднительным средствами классического аппарата корреляционного и регрессионного анализа. Более того, доказаны теоремы, согласно которым искусственная нейронная сеть типа многослойный персептрон с одним скрытым слоем способна аппроксимировать произвольную непрерывную функцию с любой степенью точности посредством изменения количества нейронов скрытого слоя.

Таким образом, использование искусственных нейронных сетей для решения задач прогнозирования электропотребления является перспективным направлением исследований благодаря способности воспроизведения сложных нелинейных зависимостей и аппроксимации сколь угодно сложных функций, а также наличию других ценных качеств (способность к обобщению, устойчивость к помехам, отсутствие ограничений на характер входной информации и др.).