

Розглянуто задачу підвищення точності прогнозування знаку і діапазону зміни добового побутового електроспоживання. Запропоновано перейти від задачі регресії до задачі класифікації. Визначено оптимальний набір категорій. На основі нейронних мереж каскадної кореляції синтезовано категоріальний предиктор електроспоживання

Ключові слова: Прогнозування, побутове електроспоживання, нейронні мережі

Rассмотрена задача повышения точности прогнозирования знака и диапазона изменения суточного бытового электропотребления. Предложено перейти от задачи регрессии к задаче классификации. Определен оптимальный набор категорий. На основе нейронных сетей каскадной корреляции синтезирован категориальный предиктор электропотребления

Ключевые слова: Прогнозирование, бытовое электропотребление, нейронные сети

The task of increase of prediction precision of daily domestic power consumption alteration sign and range is considered. It is suggested to pass from the task of regression to the task of classification. The optimum set of categories is determined. The categorical power consumption predictor based on cascade correlation neural network synthesized

Key words: Prognostication, domestic electro-consumption, neuron networks

КАТЕГОРИАЛЬНОЕ ПРОГНОЗИРОВАНИЕ СВЯЗНОГО ЭНЕРГОПОТРЕБЛЕНИЯ В КОММУНАЛЬНО- БЫТОВОМ СЕКТОРЕ

Д. И. Вороненко
Аспирант*
E-mail: dvonko@gmail.com

К. В. Махотило
Кандидат технических наук, старший научный сотрудник,
доцент*
*Кафедра электрических станций
Национальный технический университет «Харьковский
политехнический институт»
ул. Фрунзе, 21, г. Харьков, 61002
E-mail: mahotilo@yandex.ua

1. Введение

Прогнозированию электрической нагрузки регионов и энергосистем является актуальной задачей, решению которой посвящено множество зарубежных и отечественных публикаций [1, 2]. Наиболее современным и перспективным методом математического моделирования в данных исследованиях являются искусственные нейронные сети.

Общепринятым подходом к прогнозированию электропотребления на следующие сутки является создание предикторов регрессионного типа, которые напрямую прогнозируют числовую величину электрической нагрузки на завтра. Выполненный анализ точности таких предикторов, в том числе и результатов собственных исследований [3], обнаружил три существенных недостатка этого подхода.

Первым недостатком является сложность достижения и поддержания низкой ошибки прогнозирования, которая для лучших моделей составляет в среднем 3–5%. Медленные, но постоянные изменения в структуре и характере энергопотребления, ошибки учета энерго-

потребления и достаточно высокая погрешность метеопрогнозов ограничивают максимально достижимую точность прогнозирования на сутки.

Вторым недостатком является сложность перетренировки нейросетевой модели для поддержания требуемого качества прогноза. Это является автоматизированной задачей, но вычислительно высокоинтенсивной. Кроме этого, результативность перетренировки не является гарантированной и может потребовать квалифицированного вмешательства разработчиков модели.

Третий недостаток связан с особенностями практического использования прогноза электропотребления при существующем уровне развития функций АСКУЭ энергопредприятий [4]. Предиктор, генерирующий прогноз величины электропотребления, не может рассматриваться как прямое средство поддержки принятия решения диспетчерским персоналом. Его выход требует дополнительной аналитической обработки для выделения значимых ситуаций и выбора оптимальной последовательности действий, а также преобразования в удобную для восприятия человеком форму.

В то же время, анализ задач, решаемых при планировании работы и топливообеспечения региональных энергокомплексов, показывает, что с практической точки наибольшую важность зрения представляет не абсолютная точность предсказания величины спроса на энергию, а правильность предсказания знака и диапазона ожидаемого его изменения. Это позволяет совершить переход от задачи регрессии, т.е. прямого прогнозирования энергопотребления, к задаче классификации – прогнозированию категории изменения энергопотребления по отношению к текущему значению.

Этот подход лишен или в значительной мере менее подвержен указанным выше недостаткам. В частности, при нем снижается необходимость в перестройке прогностических моделей и упрощается процесс их обучения. Также прогностическая модель-классификатор может быть натренирована на выделение в динамике изменения энергопотребления ситуаций, важных для диспетчерского персонала, и ее выходы могут быть прямо преобразованы в управляющие команды. Такой подход дает нам возможность сконцентрировать усилия на повышении точности прогнозирования динамики изменения спроса на электроэнергию в коммунально-бытовом секторе. Это свою очередь позволит повысить эффективность управления региональным топливно-энергетическим комплексом.

Целью настоящей работы является развитие методов моделирования и прогнозирования электропотребления в быту, накопленного за последнее десятилетие харьковскими учеными и энергетиками [5], путем применения предикторов классификационного типа. В первой части статьи изложены методические особенности подхода, разработанного для оперативного прогнозирования суточного электропотребления крупным жилым массивом мегаполиса на основе классификационного подхода на примере г. Харькова. Вторая часть посвящена синтезу прогностической модели категориального типа и оценке ее качества.

2. Переход к задаче классификации

Схематично, суть перехода от задачи прямого прогнозирования суточного электропотребления на следующие сутки W_{k+1} к задаче классификации класса изменения в потреблении ΔW_{k+1} , представлена на рис. 1.

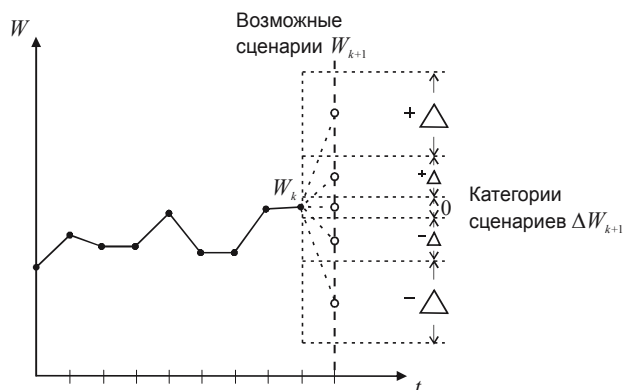


Рис. 1. Схема перехода от задачи регрессии к задаче классификации

В задаче классификации нейронная сеть должна отнести каждое наблюдение x к одному из m классов X_m или, в общем случае, оценить вероятность принадлежности наблюдения к каждому из классов $P(x \in X_m)$. В данной работе для классификации используется номинальная выходная переменная – различные ее значения соответствуют различным классам.

В такой ситуации возникает проблема, связанная с представлением данных о нескольких классах для выхода нейронной сети, решение которой традиционно достигается при помощи кодирования по методу «1-из-N».

Наиболее простым способом представления выходных данных является вектор, компоненты которого соответствуют различным номерам классов: $X = (X_1, \dots, X_m)$. При этом i -я компонента вектора X соответствует i -му классу, прочие же компоненты устанавливаются в 0: $X = (0, \dots, 0, 1, 0, \dots, 0)$. При интерпретации результата обычно считается, что номер класса определяется номером выхода сети, на котором появилось максимальное значение. Например, если для сети с m выходами y_1, \dots, y_m , максимальное значение имеет компонента $y_j = \max(y_1, \dots, y_m)$, т.е. компонента с номером j , то данное наблюдение относится к классу X_j .

Для логистической выходной функции активации наиболее подходящими значениями по умолчанию являются 0 и 1.

При кодировании «1-из-N» на каждый класс изменения энергопотребления выделяется в отдельный элемент (выходной нейрон). При наличии нескольких выходных нейронов y_1, \dots, y_m каждый из них будет содержать числовые значения $y_j \in [0, 1]$, и потребуется специальный расчет для того, чтобы по результатам работы нейросети определить выходной класс X_j .

Чтобы уверенно определить класс по набору выходных значений, нейросеть должна решить, «достаточно ли близки» они к нулю или единице. Если такой близости не наблюдается, класс считается «неопределенным».

В данном подходе, для интерпретации выходных значений y_j используются пороги принятия α и отвержения β . Определенный класс X_j выбирается только в том случае, если значение соответствующего выходного элемента выше порога принятия, $y_j > \alpha$, а всех остальных выходных элементов - ниже порога отвержения $y_j < \beta, j \neq j^*$. Если же данное условие не выполнено, класс не определяется.

Фактически пороги принятия и отвержения являются дополнительными параметрами модели, которые влияют на ее качество и установка которых требует дополнительных экспериментов и вычислительных усилий. Пороговые значения можно корректировать, чтобы заставить нейросеть быть более или, наоборот, менее «решительной» при объявлении класса.

С точки зрения задачи прогнозирования энергопотребления, отсутствие определенного класса неприемлемо, так как ставит перед системой управления или принимающим решение диспетчером, дополнительную задачу – что делать, если класс не определен. Поэтому в данной работе было решено установить порог принятия $\alpha = 0$, а порог отвержения $\beta = 1$. При $\alpha = 0$ выходной сигнал выигравшего элемента всегда будет принят, а при $\beta = 1$ все остальные элементы неизбежно будут отвергнуты, и поэтому алгоритм сводится к простому выбору выигравшего элемента.

При такой настройке, которую часто называют «победитель получает все» или “winner-takes-all” (WTA), выходной нейрон с наивысшим уровнем активации будет определять класс вне зависимости оттого, что происходит в других элементах. Более того, нейросеть классификационного типа с такой настройкой будет наиболее решительной и выдаст наиболее близкий ответ (класс) в любом случае.

Структурная схема слоя нейронов типа WTA представлена на рис. 2.

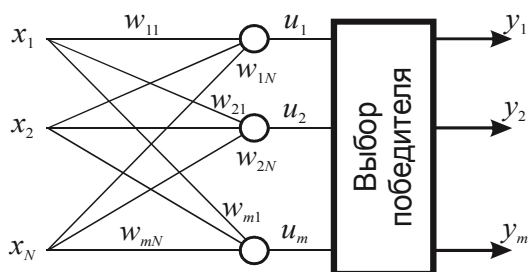


Рис. 2. Структурная схема слоя нейронов типа WTA

Каждый конкурирующий нейрон в группе получает одни и те же входные сигналы. Каждый нейрон рассчитывает выходной сигнал своего сумматора обычным образом: $u_i = \sum_j w_{ij}x_j$.

По результатам сравнения всех $u_i, i = \overline{1, m}$, выбирается нейрон-победитель, обладающий наибольшим значением $u_j = \max(u_j)$. Выходной сигнал нейрона-победителя y_j получает значение 1, выходные сигналы всех остальных нейронов – 0.

Для приведения формы выходного сигнала нейросетевой модели в форму удобную для восприятия человеком, выходной слой нейросети был снабжен тремя группами нейронов, одна из которых предназначена для вывода прогностической информации в случае положительных приростов в электропотреблении, вторая – для вывода этой информации в случае отрицательных приростов, а третья – для вывода этой информации в случае отсутствия прироста. Количество нейронов в каждой группе выбирается так, что каждый нейрон соответствует конкретному типу при-

роста спроса на электроэнергию, что напрямую идентифицирует важность прогнозируемой информации.

Дополнительно, для всех нейронов y_j выходного слоя была использована специальная функция пост-процессирования порогового типа:

$$f(y) = \begin{cases} 1, & y \geq T \\ 0, & y \in \{y_j\}. \end{cases}$$

Пороговая функция преобразует числовой сигнал нейросети y в сигнал одного из двух уровней: 0 или 1, которые интерпретируются как признак реализации в будущем того или иного типа прироста спроса.

Для каждого набора входных сигналов x будет активизирован только один нейрон выходного слоя $y_j(x)$, таким образом, каждый нейрон идентифицирует только один тип прироста спроса $X_{y_j(x)}$. В результате на выходе предиктор генерирует не непрерывные значения величины, а целочисленные признаки, которые указывают на направленность динамики будущего изменения спроса на электроэнергию.

3. Категориальный предиктор на основе нейросети каскадной корреляции

Для синтеза категориального предиктора и проведения сравнительных исследований были использованы данные о суточном потреблении электроэнергии Алексеевским жилым массивом г. Харькова, об отпуске тепла Харьковской ТЭЦ-5 и метеоданные харьковского гидрометеоцентра из архивов АПК «Зевс» [5] за несколько лет наблюдений. Набор тренировочных шаблонов был скомпонован из данных за три последовательных отопительных сезона. В качестве проверочной выборки использовались архивы за последующий отопительный сезон.

Что бы определить оптимальное количество категорий и их диапазоны, было решено провести статистический анализ распределения изменений W_k , а затем совместить его с экспертным анализом значимости различных величин изменений с точки зрения их влияния на оперативное управление энергосистемой.

На рис. 3 приведен график распределения изменений, на котором указаны значения, взятые из вышеуказанного набора данных.

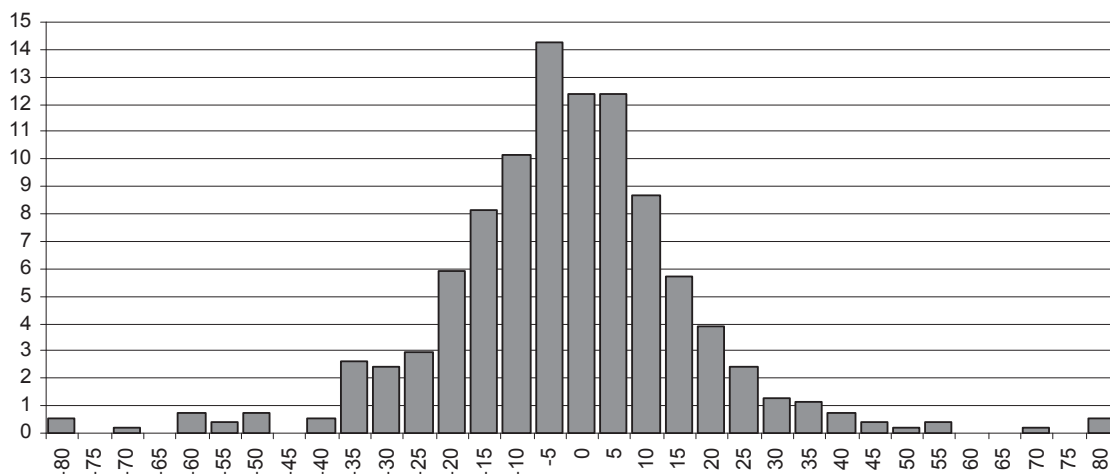


Рис. 3. График распределения суточных изменений в энергопотреблении

Проведя статистический анализ распределения изменений в энергопотреблении, а также, учитывая тот факт, что бесосновательное увеличение количества категорий приводит к снижению качества прогнозирования и усложнению соответствующих им управляющих правил, были выделены 7 категорий изменения динамики спроса: 3 положительных, 3 отрицательных и нулевая категория, т.е. незначительных изменений.

Три положительных категории изменения представляют собой: положительный прирост величиной более 40 МВт·час, положительный прирост в диапазоне от 35 МВт·час до 40 МВт·час и положительный прирост в диапазоне от 15 МВт·час до 35 МВт·час. Аналогично им выделены три отрицательные категории с симметричными границами диапазонов. В нулевую категорию отнесены все изменения в диапазоне [-15; 15] МВт·час.

Выбор порога в 15 МВт·час был сделан исходя из уровня естественных флуктуаций в энергопотреблении жилого массива при одинаковых метео- и других внешних условиях. Эта величина составляет 3% от максимального уровня суточного потребления и соответствует колебаниям среднечасовой нагрузки в пределах 600 КВт. Для исследуемого жилого массива с населением более 200 тыс. человек такие изменения нагрузки соответствуют случайным изменениям в количестве включенных осветительных приборов.

Набор входных переменных, который использовался для синтеза предиктора, представлен в табл. 1.

Таблица 1

Набор входных переменных

№	Наименование переменной	Обозначение
1	Номер дня в полугодии светового цикла	ρ_{k+1}
2	Тип дня недели	γ_{k+1}
3	Прогноз минимальной температуры на завтра	\underline{T}_{k+1}
4	Прогноз максимальной температуры на завтра	\overline{T}_{k+1}
5	Минимальная температура в текущих сутках	\underline{T}_k
6	Максимальная температура в текущих сутках	\overline{T}_k
7	Планируемый на завтра отпуск тепла	Q_{k+1}
8	Потребление тепла в текущих сутках	Q_k
9	Потребление электроэнергии в текущих сутках	W_k

Переход к задаче прогнозирования категорий приводит к необходимости поиска нейростевок архитектур, которые особо продуктивны для задач классификации. После сравнения на тестовых наборах данных эффективности многослойных перцептронов, РБФ-сетей, сетей Кохонена и нейросетей каскадной корреляции [6], был сделан выбор в пользу последних ввиду лучших результатов, получаемых в данной предметной области. Этот выбор также объясняется с теоретической точки зрения, так как каскадные нейросети используют ковариацию вместо корреляции для вычисления ошибки нейросети во время тренировки.

Целевая переменная W_k была преобразована к категориальному виду. Диапазоны для каждой категории целевой переменной ΔW_{k+1} представлены в табл. 2.

Таблица 2

Диапазоны категорий целевой переменной

Категория	H^-	h^-	η^-	o	η^+	h^+	H^+
ΔW , МВт·ч	<-40	[-35; -40]	[-15; -35]	[-15; 15]	[15; 35]	[35; 40]	>40

На первой стадии исследований для построения модели использовался тренировочный набор данных, однако ввиду относительной малочисленности данных и невозможности представить в тестовом наборе равномерно все категории с достаточным количеством записей в каждой категории, в дальнейшем анализ работы модели проводился на полном наборе данных.

Далее был проведен анализ свойств категориального предиктора на основе нейросетей каскадной корреляции. При оценке способностей нейросети к классификации используется таблица статистик классификации, в которой указывается, сколько наблюдений было классифицировано правильно. Для наилучшей модели суммарная точность составила 84,3%. Более подробно результаты работы модели представлены в табл. 3.

Таблица 3

Успешность классификации прогноза суточного электропотребления на обучающем и полном наборе данных, %

Набор	Категория						
	H^-	h^-	η^-	o	η^+	h^+	H^+
Обучающий	62,5	100	72,9	94,3	70	100	80
Полный	55,6	100	71,2	91,4	65,3	100	80

Для дополнительного подтверждения практической значимости результата было проведено сравнение полученных категориальных моделей с работой наилучшей модели прямого числового прогнозирования, описанной в предыдущей работе [3], которая напрямую прогнозирует числовое значение энергопотребления на завтра.

Сравнение проводилось в два этапа. На первом этапе результат, генерируемый числовой моделью, преобразовывался в категорию согласно диапазонам целевой переменной категориальной модели, указанным в табл. 2, и полученная в результате категория сравнивалась с реальной категорией, т.е. категорией вычисленной по той же табл. 2, но с использованием реальных (а не спрогнозированных) значений электропотребления. На втором этапе для числовой модели для каждой категории был вычислен процент верно классифицированных категорий, который был сравнен с процентом верно классифицированных категорий категориальной модели из табл. 3. Результаты сравнения приведены в табл. 4.

Таблица 4

Успешность классификации прогноза суточного электропотребления числовой и категориальной моделями, %

Модель	Категория						
	H ⁻	h ⁻	η ⁻	o	η ⁺	h ⁺	H ⁺
Числовая	60	0	38,9	64,8	18,5	0	22,2
Категориальная	85,7	80	64,3	90,5	72,2	100	80

4. Выводы

При планировании режимов работы предприятий ТЭК большую практическую ценность представляет не абсолютная точность предсказания величины ожидаемого спроса, а правильность предсказания знака и диапазона ожидаемого изменения суточного потребления. Этот показатель можно существенно улучшить путем перехода от задачи регрессии (прогнозирование числового значения изменения) к задаче классификации (прогнозирование категории изменения).

Выполненный в работе анализ электропотребления жилого массива позволил определить оптимальное количество категорий, их числовые границы, а также синтезировать для него на основе нейронных сетей каскадной корреляции категориальный предиктор.

Сравнение точности категориального предиктора с ранее полученными моделями числового прогнозирования доказал превосходство нового подхода с точки зрения точности и практической применимости прогнозов.

Литература

1. Меламед А. Современные методы анализа и прогнозирования режимов электропотребления в электроэнергетических системах / А. М. Меламед // В кн.: Прогнозирование и управление электропотреблением в электроэнергетических системах. Сер. «Энергетические системы и их автоматизация» (Итоги науки и техники) / А. М. Меламед – М.: ВИНТИ, 1988. – Т.4. – С.4–111.
2. Бэнн Д. В., Сравнительные модели прогнозирования электрической нагрузки / Д. В. Бэнн, Е. Д. Фармер. – М.: Энергоатомиздат., 1987. – 200 с.
3. Вороненко Д. В. Опыт синтеза и применения предикторов на базе искусственных нейронных сетей для прогнозирования спроса на электроэнергию в бытовом секторе / Д. В. Вороненко, Г. К. Вороновский // Информационные технологии в энергетике : труды 3-й научно-практической конференции, 2004 г.– Киев, ИПМЭ, 2004. – С.45–51.
4. Автоматизация диспетчерского управления в электроэнергетике / [Баринов В. А., Гамм А. З. и др.] – М.: МЭИ, 2000. – 648 с.
5. Вороновский Г.К. Усовершенствование практики оперативного управления крупными теплофикационными системами в новых экономических условиях / Геннадій Кирилович Вороновський. – Х.: Изд-во «Харьков», 2002. – 240с.
6. Использование каскадных нейросетей Фалмана для прогнозирования связанного потребления электричества и тепла / Д. В. Вороненко // Вісник Національного технічного університету «Харківський політехнічний інститут». Збірник наукових праць. Тематичний збірник наукових праць «Проблеми автоматизованого електроприводу Теорія і практика», – Харків: НТУ"ХПІ", 2004 – Т. 1, №43. – С. 93–97.