

Автоматизация. Современные Технологии

ЕЖЕМЕСЯЧНЫЙ МЕЖОТРАСЛЕВОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Главный редактор

Шахнов В.А. — д.т.н., проф., член-кор. РАН,
МГТУ им. Н.Э. Баумана

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Афанасьев В.Н. — д.т.н., проф., МИЭМ НИУ ВШЭ
Басараб М.А. — д.ф.-м.н., проф., МГТУ им. Н.Э. Баумана
Большаков А.А. — д.т.н., проф., СПбГТИ (ТУ)
Булдакова Т.И. — д.т.н., проф., МГТУ им. Н.Э. Баумана
Ван Мэйлин — д.т.н., проф., Пекинский
политехн. ун-т (КНР)
Зинченко Л.А. — д.т.н., проф., МГТУ им. Н.Э. Баумана
Зубов Н.Е. — д.т.н., проф., РКК «Энергия»
Кларк Р. — д.т.н., проф., КИУ (г. Ворвик,
Великобритания)
Криони Н.К. — д.т.н., проф., УГАТУ (г. Уфа)
Кузнецов А.Е. — д.т.н., проф., РГРТУ (г. Рязань)
Мальцева С.В. — д.т.н., проф., НИУ ВШЭ
Микаева С.А. — д.т.н., проф., МГУПИ
Неусыпин К.А. — д.т.н., проф., МГТУ им. Н.Э. Баумана
Нефёдов Е.И. — д.ф.-м.н., ИРЭ РАН
Никифоров В.М. — д.т.н., проф., ФГУП «НППАП
им. Н.А. Пилюгина»
Пролетарский А.В. — д.т.н., проф., МГТУ им. Н.Э. Баумана
Проталинский О.М. — д.т.н., проф., НИУ МЭИ
Путилов В.Н. — ООО «Изд-во "Инновационное
машиностроение"» (заместитель
главного редактора)
Румянцева О.Н. — генеральный директор ООО «Изд-во
"Инновационное машиностроение"»
Ся Юаньцин — д.т.н., проф., Пекинский
политехн. ун-т (КНР)
Фу Ли — д.т.н., проф., Ин-т Бейхан (КНР)
Фёдоров И.Б. — д.т.н., проф., академик РАН,
президент МГТУ им. Н.Э. Баумана
Хэ Юн — д.т.н., проф., Нанкинский ун-т
науки и технологий (КНР)
Чистякова Т.Б. — д.т.н., проф., СПбГТИ (ТУ)
Шибанов Г.П. — д.т.н., проф., Гос. лётно-испытат.
центр им. В.П. Чкалова

Редакторы — **Мымрина И.Н., Селихова Е.А.**
Компьютерная вёрстка — **Конова Е.В.**

Адрес редакции:

107076, Москва, Колодезный пер., д. 2а, стр. 2.
Тел.: 8 (499) 268-41-77.
E-mail: ast@mashin.ru; astmashin@yandex.ru;
http://www.mashin.ru

Учредитель:

ООО «НАУЧНО-ТЕХНИЧЕСКОЕ ИЗДАТЕЛЬСТВО
"ИННОВАЦИОННОЕ МАШИНОСТРОЕНИЕ"»

Журнал зарегистрирован в Роскомнадзоре 29 мая 2014 г.
(ПИ № ФС77-58102), входит в перечень утверждённых ВАК
при Минобрнауки России изданий для публикации трудов
соискателей учёных степеней, а также в систему Россий-
ского индекса научного цитирования (РИНЦ)

Издаётся с 1947 г.

Том 74

9
2020

СОДЕРЖАНИЕ

АВТОМАТИЗАЦИЯ НАУЧНО-ИССЛЕДОВАТЕЛЬСКИХ И ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ

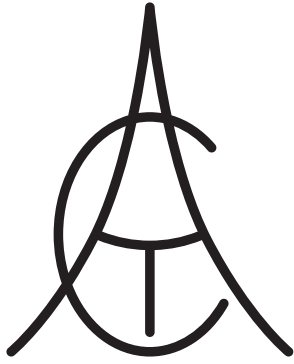
- Лукьяненко В.С., Щеголев А.В., Сергиенко С.Н.** Автомати-
зация гальванического цинкования стальных изделий 387
Жеглова Ю.Г., Титаренко Б.П. Анализ проектных решений
ограждений котлованов 390
**Кабалдин Ю.Г., Шатагин Д.А., Кузьмишина А.М., Ано-
сов М.С., Колчин П.В.** Цифровой двойник оборудования как
основа для потребителя в условиях цифрового производ-
ства 394
Данилов К.В. Прогнозирование энергопотребления на
основе автоматического машинного обучения 402

СОВРЕМЕННЫЕ ТЕХНОЛОГИИ

- Чжан Лифей.** Улучшение характеристик нелинейных моде-
лей динамических систем 408
Железникова О.Е., Мышонков А.Б., Микаева С.А. Иссле-
дования режимов облучения листового салата 413
Рублева Д.Б., Смирнов Н.Я. Исследование устойчивости
алгоритмов распознавания образов к внешним воздей-
ствиям 418

СИСТЕМЫ И ПРИБОРЫ УПРАВЛЕНИЯ

- Васильев М.И., Васильев И.И.** Поиск оптимальной посто-
янной скорости изменения регулирующего входного воздей-
ствия нечёткого регулятора стабилизирующего контура . . . 424
Микаева А.С., Микаева С.А. Контроллер управления
NF-K-001, NF-K-002 428



Automation. Modern Technologies

MONTHLY INTERBRANCH SCIENTIFIC AND TECHNICAL JOURNAL

Chief in Editor

Shakhnov V.A. – dr.en.s., prof., corr. member of RAS,
MSTU behalf of N.E. Bauman

EDITORIAL COUNCIL:

Afanasyev V.N. – dr.en.s., prof., MIEM NRU HSE

Basarab M.A. – dr.ph.-m.s., prof., MSTU behalf
of N.E. Bauman

Bolshakov A.A. – dr.en.s., prof., SPbSIT (TU)

Buldakova T.I. – dr.en.s., prof., MSTU behalf
of N.E. Bauman

Wang Meiling – Ph.D., prof., Beijing Ins. of Technology,
P.R. China

Zinchenko L.A. – dr.en.s., prof., MSTU behalf
of N.E. Bauman

Zubov N.E. – dr.en.s., prof., RSC «Energy»

Clark R. – Ph.D., prof., EMB of Warwick Un., UK

Krioni N.K. – dr.en.s., prof., USATU, Ufa

Kuznetsov A.E. – dr.en.s., prof., RSREU, Ryazan

Maltseva S.V. – dr.en.s., prof., NRU HSE

Mikaeva S.A. – dr.en.s., prof., MSUPI

Neusypin K.A. – dr.en.s., prof., MSTU behalf
of N.E. Bauman

Nefedov E.I. – dr.ph.-m.s., IRE RAS

Nikiforov V.M. – dr.en.s., prof., FSUE «NPCAP
behalf of N.A. Pilyugin»

Proletarskiy A.V. – dr.en.s., prof., MSTU behalf
of N.E. Bauman

Protalinsky O.M. – dr.en.s., prof., NRU MPEI

Putilov V.N. – LLC «"Innovative Mashinostroenie"
Publishers» (deputy chief editor)

Rumyantseva O.N. – General Director of LLC «"Innovative
Mashinostroenie" Publishers»

Xia Yuantsin – Ph.D., prof., Beijing Ins. of Technology,
P.R. China

Fu Li – Ph.D., prof., Beikhan Un., P.R. China

Fedorov I.B. – dr.en.s., prof., academician RAS, President
of MSTU behalf of N.E. Bauman

He Yung – Ph.D., prof., Nanjing Un. of Science
and Technology, P.R. China

Chistyakova T.B. – dr.en.s., prof., SPbSIT (TU)

Shibanov G.P. – dr.en.s., prof., State Flight Test Center
behalf of V.P. Chkalov

Editors — **Mymrina I.N., Selikhova E.A.**

Computer layout — **Konova E.V.**

Editorial address:

107076, Moscow, Kolodezny lane — 2a, build. 2

Tel.: +7 (499) 268-41-77.

E-mail: ast@mashin.ru; astmashin@yandex.ru;

<http://www.mashin.ru>

Founder:

LLC «"INNOVATIVE MASHINOSTROENIE" PUBLISHERS»

The journal was registered in the Roskomnadzor on May 29, 2014 (PI No. FS77-58102), it is included in the list of publications approved by the Higher Attestation Commission (VAK) of the Russian Ministry of education and science for publication of the works of applicants for academic degrees and the system of the Russian index scientific citation (RINC)

It is published since 1947

Volume 74 $\frac{9}{2020}$

CONTENTS

AUTOMATION OF SCIENTIFIC-RESEARCH AND PRODUCTION PROCESSES

Lukyanenko V.S., Shchegolev A.V., Sergienko S.N. Automation
of the steel products galvanic zincing 387

Zheglava Yu.G., Titarenko B.P. Analysis of design solutions for
foundation pit fencing 390

**Kabaldin Yu.G., Shatagin D.A., Kuzmishina A.M., Anosov M.S.,
Kolchin P.V.** Digital twin of equipment as a basis for the consumer
in digital production 394

Danilov K.V. Energy prediction based on automatic machine
teaching 402

MODERN TECHNOLOGIES

Chzhan Lifye. Improving the nonlinear models characteristics of dy-
namical systems 408

Zheleznikova O.E., Myshonkov A.B., Mikaeva S.A. Studies of
irradiation regimes of leaf lettuce 413

Rubleva D.B., Smirnov N.Ya. The stability investigation of pattern
recognition algorithms to external impacts 418

SYSTEMS AND CONTROL DEVICES

Vasilyev M.I., Vasilyev I.I. Search for the optimal constant rate of
change for the adjusting input impact of the stabilizing circuit fuzzy
regulator 424

Mikaeva A.S., Mikaeva S.A. Control controller NF-K-001, -002 428

8. https://en.wikipedia.org/wiki/Bootstrap_aggregating (дата обращения: 15.12.2018).

9. [https://en.wikipedia.org/wiki/Boosting_\(machine_learning\)](https://en.wikipedia.org/wiki/Boosting_(machine_learning)) (дата обращения: 15.12.2018).

10. **Лоладзе Т.Н.** Прочность и износостойкость режущего инструмента. М.: Машиностроение, 1982. 320 с.

11. **Григоревич В.К.** Металлическая связь и структура металлов. М.: Наука, 1988. 296 с.

12. **Андриевский А.Р., Спивак И.И.** Прочность тугоплавких соединений и материалов на их основе. Справочник. Челябинск: Металлургия, 1989. 368 с.

13. **Самсонов Г.В., Винницкий И.М.** Тугоплавкие соединения (справочник). М.: Металлургия, 1976. 560 с.

14. **Верещака А.С.** Работоспособность режущего инструмента с покрытием. М.: Машиностроение, 1993. 336 с.

15. **Выбор** состава и структуры износостойких наноструктурных покрытий для твердосплавного режущего инструмента на основе квантово-механического моделирования / Ю.Г. Кабалдин, О.В. Кретицин, Д.А. Шатагин, Е.Е. Власов. М.: Инновационное машиностроение, 2017. 216 с.

Ссылка для цитирования

Кабалдин Ю.Г., Шатагин Д.А., Кузьмишина А.М., Аносов М.С., Колчин П.В. Цифровой двойник оборудования как основа для потребителя в условиях цифрового производства // Автоматизация. Современные технологии. 2020. Т. 74. № 9. С. 394–402. DOI: 10.36652/0869-4931-2020-74-9-394-402.

УДК 004.8:519.688

К.В. Данилов

(ПАО «НЛМК», Национальный исследовательский университет «Высшая школа экономики», г. Москва)

danilovkostya@yandex.ru

ПРОГНОЗИРОВАНИЕ ЭНЕРГОПОТРЕБЛЕНИЯ НА ОСНОВЕ АВТОМАТИЧЕСКОГО МАШИННОГО ОБУЧЕНИЯ

Рассмотрена задача прогнозирования энергопотребления на основе автоматического машинного обучения. Приведена схема процесса автоматического создания и применения модели прогнозирования. Предлагаемый подход апробирован на основе данных о потреблении электроэнергии в регионах России. Проведённый вычислительный эксперимент показал высокую эффективность разработанной модели. Точность прогнозирования составила 97...99 %.

Ключевые слова: автоматическое машинное обучение; прогнозирование временных рядов; прогнозирование потребления электроэнергии; деревья решений; рекуррентные нейронные сети.

The problem of energy consumption prediction based on automatic machine teaching is considered. The scheme of the automatic creation process and application of the prediction model is given. On the basis of electricity consumption data in the regions of Russia the proposed approach is tested. A carried out computational experiment showed the high efficiency of the developed model. The prediction accuracy is 97...99 %.

Keywords: automatic machine teaching; time series prediction; prediction of the electricity consumption; decision trees; recurrent neural networks.

Введение. Прогнозирование объёмов потребления электроэнергии является актуальной задачей, так как напрямую влияет на затраты компаний. Точный прогноз позволяет нивелировать расходы и добиться максимальной прибыли. Безусловно, для решения подобной задачи требуются данные (историческая выборка), которые отображают энергопотребление в прошлые периоды. Как правило, такие данные представлены в виде временных рядов, анализ которых позволяет построить прогноз.

Временной ряд — дискретная последовательность чисел, характеризующая состояние объекта наблюдения в отдельные периоды времени [1]. Существует большое количество методов прогнозирования временных рядов, каждый из которых имеет свои преимущества и недостатки. В данной работе используется автоматическое машинное обучение для прогнозирования объёмов энергопотребления.

Автоматическое машинное обучение — подход к построению модели машинного об-

учения, основанный на сквозном автоматическом процессе предварительной обработки данных, генерации и выбора дополнительных признаков, обучении и тестировании модели.

Новизна предлагаемого подхода заключается в использовании комбинации нейросетевых технологий и деревьев решений на основе автоматического машинного обучения, что предоставляет следующие преимущества: повышение точности прогнозирования временного ряда; оптимизация гиперпараметров модели в автоматическом режиме; автоматическое переобучение модели для заданного периода; автоматический выбор наиболее оптимального набора признаков для обучения/переобучения модели; проведение валидации моделей в автоматическом режиме.

Предлагаемый подход к созданию модели прогнозирования объёмов потребления электроэнергии позволяет не только оптимизировать временные затраты, но и улучшить качество модели.

Постановка задачи. Пусть имеется временной ряд $Y = \{Y_1, \dots, Y_i\}$, где Y_i — i -е значение элемента временного ряда, $i, Y_i \in R$. Необходимо определить значения элементов временного ряда в последующие моменты времени $\{Y_{i+1}, Y_{i+2}, \dots, Y_{i+n}\}$, где величина $n \in R$ устанавливает глубину прогнозирования: краткосрочное (1...7 дней), среднесрочное (8...180 дней), долгосрочное (более 180 дней) [2]. Важно отметить, что диапазоны глубины прогнозирования зачастую зависят от предметной области задачи.

В данной работе точность прогнозирования временного ряда оценивается с помощью известных показателей [3]:

$MAPE$ (mean absolute percentage error) — среднее абсолютное отклонение в процентах,

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i - Y'_i|}{Y_i}; \quad (1)$$

$RMSE$ (root mean squared error) — среднеквадратичное отклонение,

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - Y'_i)^2}; \quad (2)$$

MAE (mean absolute error) — среднее абсолютное отклонение,

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - Y'_i|, \quad (3)$$

где Y_i — фактическое значение элемента временного ряда; Y'_i — прогнозное значение элемента временного ряда; N — число элементов во временном ряду.

Основные методы и подходы к решению задачи. Методы прогнозирования можно разделить на качественные и количественные. В большинстве случаев методы *качественного прогнозирования* требуют заключений экспертов предметной области и используются в тех случаях, когда выборка исторических данных крайне мала. Наиболее известными качественными методами прогнозирования являются методы Дельфи, Кука, социологических исследований и прогноза по аналогии [10].

Методы *количественного прогнозирования* опираются на накопленную историческую информацию, которая предоставляет возможность обучить математическую модель описывать закономерности и паттерны, имеющиеся в данных. Количественный подход к прогнозированию временных рядов на основе ARIMA-технологии показывает удовлетворительную эффективность [6, 8, 16]. В качестве недостатка данного подхода можно отметить необходимость генерации и выбора наиболее важных признаков, которые участвуют в прогнозировании. Кроме того, перед использованием данного метода требуется добиться стационарности временного ряда, проведя итеративное исследование автокорреляционной и частично автокорреляционной функций.

Применение искусственных нейронных сетей (ИНС) для прогнозирования временных рядов широко распространено. Рекуррентные ИНС с модулями долгой краткосрочной памяти (LSTM) могут использовать свои внутренние ресурсы для обработки последовательностей произвольной длины, что позволяет делать более точные прогнозы [1, 7, 9]. Также искусственные нейронные сети обеспечивают учёт нелинейности в данных, однако для их обучения требуется большой объём исторической информации.

Гибридный подход, основанный на комбинации ARIMA-моделей и искусственных нейронных сетей [4], обладает преимуществами линейных и нелинейных моделей, что позволяет повысить точность и стабильность прогноза.

Деревья решений предоставляют возможность работы с пропущенными значениями [5, 9]. Простота интерпретации результатов и обоснования полученных прогнозов является их преимуществом. Однако необходимо контролировать размер дерева, так как высока вероятность переобучения модели. Тем не менее результаты работы [5] доказывают высокую эффективность предложенного подхода. В табл. 1 представлен сравнительный анализ

Таблица 1
Сравнительный анализ ИНС (LSTM) и ГБДР

Свойство	ИНС (LSTM)	ГБДР
Потребность в нормализации данных	+	–
Работа с пропущенными значениями	–	+
Высокая склонность к переобучению	+	–
Робастность к выбросам	–	+
Сглаженный прогноз	+	–

свойств ИНС типа LSTM и градиентного бустинга на основе деревьев решений (ГБДР).

Предлагаемый в данной работе подход основывается на комбинации искусственной нейронной сети и градиентного бустинга деревьев решений. Процесс автоматического машинного обучения позволяет более оптимально сгенерировать и отобрать дополнительные признаки для обучения и тестирования модели прогнозирования, провести настройку её гиперпараметров.

Процесс автоматического машинного обучения. Процесс автоматического машинного обучения (рис. 1) включает в себя следующие этапы: подготовка данных, генерация признаков, выбор признаков, создание модели прогнозирования временного ряда, настройка гиперпараметров модели, валидация модели.

Входные данные (временные ряды) поступают в структурированном виде с размеченными признаками и целевой переменной. Пакет данных подаётся на блок подготовки, где решается задача детектирования и обработки выбросов (пропущенных значений). Выбросы — экстремальные значения, выбивающиеся из общей выборки данных [12]. Этап подготовки данных может завершить процесс автоматического машинного обучения, если будут обнаружены недопустимые пробелы или большое количество выбросов. Генерация признаков осуществляется на предварительно

обработанных данных, в которых отсутствуют выбросы (пропущенные значения). Структуры поступающей информации на этапе подготовки данных и создания дополнительных признаков одинаковы. Блок генерации признаков содержит установленные правила формирования переменных. Далее осуществляется поиск наиболее важных признаков для решения поставленной задачи. Функция выбора признаков принимает на входе два параметра: весовой коэффициент, позволяющий отсеивать признаки по степени важности, и пакет данных, содержащий подготовленную информацию в виде временных рядов. На выходе функции выбора признаков формируется список с необходимыми параметрами для построения модели прогнозирования. Функция создания модели и настройки гиперпараметров проводит анализ алгоритмов машинного обучения и выбор наиболее оптимального для повышения качества прогноза. На этапе валидации модели проводится дополнительная верификация модели прогнозирования на отложенной выборке данных.

Предварительная обработка данных. Процесс предварительной обработки данных является одним из наиболее важных в автоматическом машинном обучении, так как вносит корректировки в поступающую информацию (рис. 2). Данный этап включает в себя поиск и обработку выбросов (пропущенных значений).

Поиск выбросов осуществляется на основе статистической оценки z-score. Выбор метода поиска выбросов определялся эмпирически. Найденные выбросы удаляют и заменяют на медианные значения скользящего окна [12, 13].

Детектирование пропущенных значений основано на допущении, что если временной ряд содержит значения, которые невозможно преобразовать в числовой формат, то их считают пропуском. В свою очередь, пропущенные значения также заменяют на медиану скользящего окна.

Результатом предварительной обработки данных может быть два различных исхода. Если есть возможность сделать прогноз на ос-

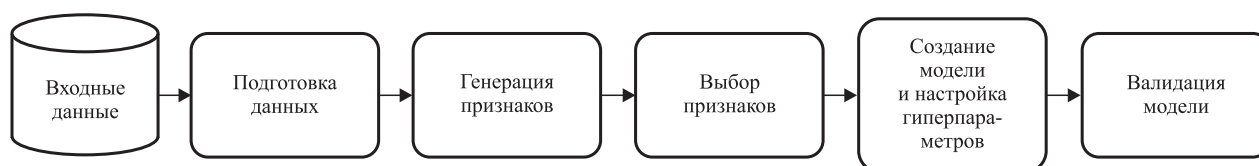


Рис. 1. Процесс автоматического машинного обучения

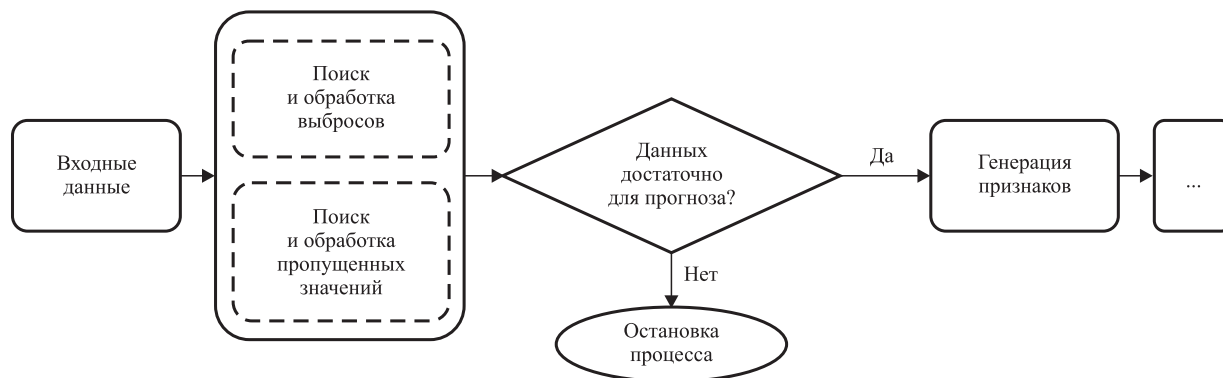


Рис. 2. Предварительная обработка данных

нове имеющихся данных, то переходят к этапу генерирования дополнительных признаков, иначе процесс автоматического машинного обучения останавливается.

Генерирование дополнительных признаков. Создание дополнительных признаков осуществлялось на основе предметной экспертизы и результатов первичного анализа данных. В частности, при обогащении данных применялись следующие подходы [9, 13]: взятие лаговых значений; тригонометрические функции; перемножение бинарных признаков; дифференцирование лаговых значений; взятие сезонной компонентны временного ряда; взятие трендовой компонентны временного ряда; скользящее окно; взятие лаговых значений по сезонной и трендовой компонентам временного ряда.

В результате процедуры генерации набор данных расширялся до 150 признаков для обучения и тестирования моделей, что позволяло повысить качество прогноза.

Выбор признаков. Протестировано два подхода при выборе наиболее важных признаков. Изначально использовался коэффициент Пирсона [6] для исследования взаимосвязи параме-

тров, но позже данный подход был отклонён, так как он не позволял учесть нелинейные зависимости. Выбор признаков для обучения и тестирования моделей осуществлялся на основе градиентного бустинга деревьев решений (ГБДР) по схеме, представленной на рис. 3.

При выборе признаков указывался порог θ , который определяли на основе множества проведённых экспериментов. В результате отбираются признаки, весовой коэффициент которых превышает заданный порог. На выходе функции выбора признаков формируется массив данных, которые участвуют в построении и валидации модели прогнозирования временного ряда.

Создание модели и выбор гиперпараметров.

На этапе создания модели применялись два подхода, основанные на искусственных нейронных сетях типа LSTM и ГБДР, реализованного в библиотеке *Xgboost* [11]. LSTM (Long short-term memory) — разновидность архитектуры рекуррентных нейронных сетей, предложенная в 1997 г. Хохрайтером и Шмидхубером. *XGBoost* (EXtreme Gradient Boosting) — библиотека машинного обучения, реализующая модель градиентного бустинга, представляющего альтернативу регрессионным методам и нейронным сетям. Метод заключается в создании ансамбля последовательно уточняющих друг друга деревьев решений [11].

Гиперпараметры — настраиваемые параметры модели перед её обучением. Оптимизация гиперпараметров проводилась перебором на основе установленной матрицы значений.

Рассмотрены четыре гиперпараметра настройки ГБДР [11, 13]:

subsample — доля объектов обучающей выборки, используемых на каждой итерации;

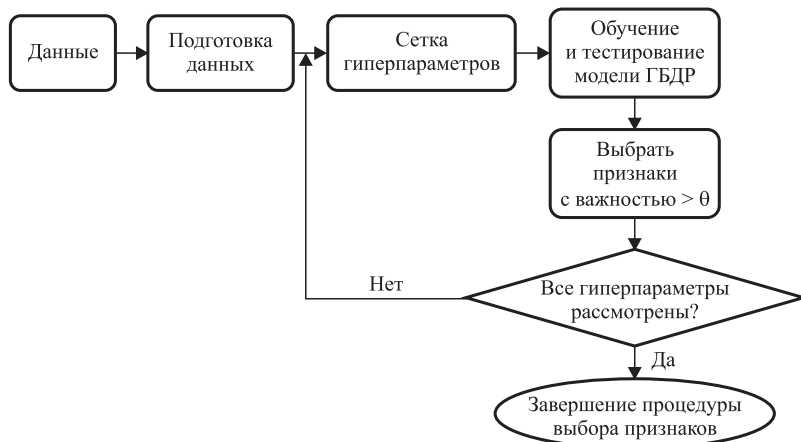


Рис. 3. Выбор признаков

colsample_bytree — доля переменных, используемых на каждой итерации;

eta — скорость обучения ансамбля моделей (критически важный параметр, контролирующий, с каким весом предсказания каждой следующей модели суммируются с предсказаниями ансамбля);

max_depth — максимальная глубина дерева (выбор слишком глубоких деревьев приводит к переобучению, а слишком маленькие деревья не позволяют эффективно восстановить искомую зависимость).

Далее рассмотрены три гиперпараметра настройки LSTM [14, 15]:

lstm_units — число lstm-юнитов в одном слое нейронной сети;

num_epochs — число эпох при обучении модели (за одну эпоху все данные в обучающей и тестовой выборках проходят через нейронную сеть);

batch_size — размер пакета данных (после прохождения всех объектов, содержащихся в одном пакете данных, следует шаг градиентного спуска для минимизации ошибки).

Результатом этапа создания модели и оптимизации её гиперпараметров является обученная модель машинного обучения, продемонстрировавшая наиболее качественный прогноз на отложенной выборке. При этом использовалась среднеквадратическая ошибка как метрика оценки качества прогноза.

Валидация модели. Этап валидации модели играет важную роль в процессе автоматического машинного обучения, так как позволяет идентифицировать переобучение модели прогнозирования и верифицировать её предсказательную способность.

Валидация модели прогнозирования проводилась на отложенной выборке данных, которая формировалась еженедельно и пополнялась вновь поступившими данными. Правила формирования отложенной выборки настраивались и устанавливались вруч-

Таблица 2

Результаты прогнозирования

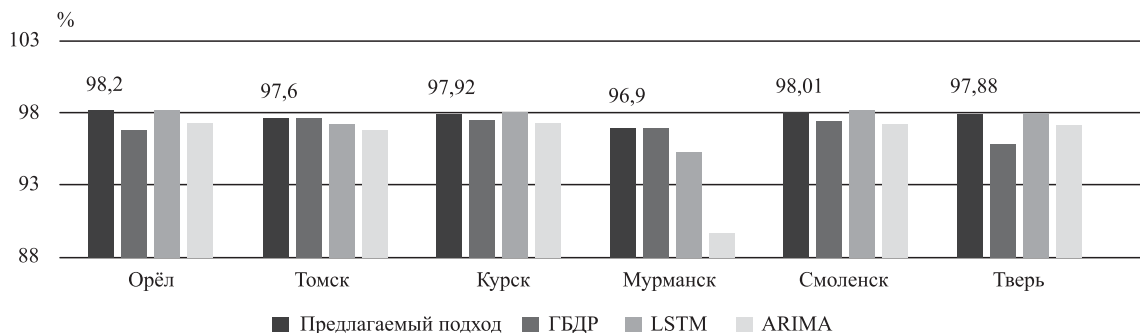
Город	Средние отклонения прогноза от фактического потребления энергопотребления, %			
	Предлагаемый подход	ГБДР	LSTM	ARIMA
Орёл	1,80	3,2	1,80	2,7
Томск	2,40	2,4	2,80	3,2
Курск	2,08	2,5	2,0	2,7
Мурманск	3,10	3,1	4,70	10,3
Смоленск	1,99	2,6	1,80	2,8
Тверь	2,12	4,2	2,05	2,9

ную на основе множества проведённых экспериментов.

Результаты экспериментов. Тестирование разработанного подхода на основе автоматического машинного обучения и ГБДР, LSTM-, ARIMA-моделей происходило на основе данных об энергопотреблении в Орле, Томске, Курске, Мурманске, Смоленске и Твери. Период апробации длился два месяца (октябрь, ноябрь) с ежедневным поступлением новых данных. Результаты прогнозирования представлены в табл. 2.

Каждый тестируемый регион имеет свою специфику и обусловленные ей допущения на правила формирования дополнительных признаков. Например, фактические значения потребления электроэнергии в Мурманске становятся доступными только в конце месяца, что ограничивает использование лаговых переменных. Влияние таких допущений отражалось негативно на точности прогнозирования. Результаты представлены ниже.

На рис. 4 представлен сравнительный анализ описываемого в данной работе метода

**Рис. 4. Точность прогнозирования (MAPE)**

и моделей на основе ГБДР, LSTM, ARIMA. Предлагаемый подход демонстрирует свою эффективность и повышенную точность прогноза. Примечательно, что при использовании комбинации нейросетевых технологий и ансамбля деревьев решений на основе автоматического машинного обучения наблюдается стабильность качественного прогноза в экстремумах временных рядов (пиках потребления электроэнергии).

Заключение. В ходе экспериментов подход к прогнозированию объёмов потребления электроэнергии с использованием комбинации нейросетевых технологий и градиентного бустинга деревьев решений на основе автоматического машинного обучения продемонстрировал точность в пределах 97...99 %. Тестирование проводилось на данных об энергопотреблении в разных городах России: Орле, Томске, Курске, Мурманске, Смоленске и Твери. Глубина исторических данных составляла 3...5 лет, в зависимости от региона. Период апробации описанного в данной работе подхода длился два месяца.

В дальнейшем планируется провести верификацию разработанного подхода на других временных рядах. Модификация этапа предварительной обработки данных и генерации дополнительных признаков имеет большой потенциал для повышения качества прогноза. Планируются внедрение новых предиктивных алгоритмов и развитие блока настройки гиперпараметров модели прогнозирования на основе байесовской оптимизации.

Библиографические ссылки

1. **Безручко Б.П., Смирнов Д.А.** Статистическое моделирование по временным рядам: учеб.-метод. пособие. Саратов: ГосУНЦ «Колледж», 2000. С. 6.
2. **Al Khafaf N., Jalili M., Sokolowski P.** Application of Deep Learning Long Short-Term Memory in Energy Demand Forecasting // 20th International Conference: Engineering Applications of Neural Networks. Communications in Computer and Information Science // Springer. 2019. Vol. 1000. P. 31.
3. **Щербаков М.В., Бребельс А., Щербакова Н.Л., Тюков А.П.** Обзор оценок качества моделей прогнозирования. URL: http://www.mtas.ru/bitrix/components/bitrix/forum.interface/show_file.php?fid=6450 (дата обращения: 05.07.2019).

4. **Büyüksahin Ü.Ç., Ertekin S.** Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition // Neurocomputing journal. 2019. Vol. 361. P. 151.

5. **Moore P.J., Lyons T.J., Gallacher J.** Random forest prediction of Alzheimer's disease using pairwise selection from time series data // PLOS ONE journal. 2018. Vol. 14. No. 2. P. 14.

6. **Feigelson E.D., Babu G.J., Caceres G.A.** Autoregressive times series methods for time domain astronomy // Frontiers in Physics journal. 2019. Vol. 6. P. 80.

7. **Waheeb W., Ghazali R.** Multi-step time series forecasting using ridge polynomial neural network with error-output feed-backs. Soft Computing in Data Science / Second International Conference. SCDS 2016. Kuala Lumpur, Malaysia. September 21—22. 2016.

8. **Shang H.L.** A robust functional time series forecasting method // Journal of Statistical Computation and Simulation. 2019. Vol. 89. No. 5. P. 795.

9. **Крючин О.В., Козадаев А.С., Дудаков В.П.** Прогнозирование временных рядов с помощью искусственных нейронных сетей и регрессионных моделей на примере прогнозирования котировок валютных пар // Исследовано в России. 2010. Т. 13. С. 354.

10. **Hallowell M.R., Gambatese J.A.** Qualitative Research: Application of the Delphi method to CEM Research // Journal of Construction Engineering and Management. 2010. Vol. 136. Iss. 1. P. 38.

11. **Jason B.** A Gentle Introduction to XGBoost for applied machine learning. URL: <https://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/> (дата обращения: 26.06.2019).

12. **Aggarwal C.C., Yu P.S.** Outlier Detection for High Dimensional Data // ACM SIGMOD Record journal. 2010. Vol. 30. No. 2. P. 10.

13. **Сидоров С.Г.** Анализ временных рядов как метод построения прогноза потребления электроэнергии // Вестник ИГЭУ. 2010. Т. 3. С. 81.

14. **Bengio Y., Frasconi P., Simard P.** The problem of learning long-term dependencies in recurrent networks / IEEE International Conference on Neural Networks. 1993. P. 1183—1188.

15. **Makridakis S., Spiliotis E., Assimakopoulos V.** Statistical and machine learning forecasting methods: Concerns and ways forward // PLOS ONE journal. 2018. Vol. 13. No. 13. P. 1—26.

16. **Conejo A.J., Plazas M.A., Espinola R., Molina A.B.** Day-ahead electricity price forecasting using the wavelet transform and ARIMA models // IEEE Transaction on Power System journal. 2005. Vol. 20. No. 2. P. 1035.

Ссылка для цитирования

Данилов К.В. Прогнозирование энергопотребления на основе автоматического машинного обучения // Автоматизация. Современные технологии. 2020. Т. 74. № 9. С. 402—407.

Уважаемые читатели!

Подписку можно оформить в любом почтовом отделении по каталогу:
«Пресса России» — индекс **27838**.