

Полтавский Артем Владимирович – Южный федеральный университет; e-mail: poltavsky@sfedu.ru; г. Ростов-на-Дону, Россия; Международный исследовательский институт интеллектуальных материалов ЮФУ; аспирант.

Полянниченко Дмитрий Сергеевич – Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте АО «НИИАС», Ростовский филиал; e-mail: d.polyanichenko@vniias.ru; г. Ростов-на-Дону, Россия; главный специалист.

Коломенская Екатерина Романовна – Южный федеральный университет; e-mail: kolomenskaya@sfedu.ru; г. Ростов-на-Дону, Россия; Международный исследовательский институт интеллектуальных материалов ЮФУ; инженер.

Бутакова Мария Александровна – Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте АО «НИИАС», Ростовский филиал; e-mail: m.butakova@vniias.ru; г. Ростов-на-Дону, Россия; д.т.н.; профессор; г.н.с.

Poltavskiy Artem Vladimirovich – Southern Federal University; e-mail: poltavsky@sfedu.ru; Rostov-on-Don, Russia; the International Research Institute of Intelligent Materials of the Southern Federal University; post-graduate student.

Kolomenskaya Ekaterina Romanovna – Southern Federal University; e-mail: kolomenskaya@sfedu.ru; Rostov-on-Don, Russia; The International Research Institute of Intelligent Materials of the Southern Federal University; engineer.

Polyanichenko Dmitry Sergeevich – JSC NIIAS; e-mail: d.polyanichenko @vniias.ru; Rostov-on-Don, Russia; team lead.

Butakova Maria Aleksandrovna – JSC NIIAS; e-mail: m.butakova@vniias.ru; Rostov-on-Don, Russia; dr. of eng. sc.; professor; chief researcher.

УДК 004.942, 654.924.56

DOI 10.18522/2311-3103-2025-1-65-81

С. Сингх, А.В. Прибыльский, Е.Ю. Косенко

РАЗРАБОТКА И ИССЛЕДОВАНИЕ АЛГОРИТМОВ ПРОГНОЗИРОВАНИЯ ПОЖАРООПАСНЫХ СИТУАЦИЙ

Раннее обнаружение пожароопасных ситуаций является критически важным аспектом обеспечения безопасности, так как позволяет минимизировать риск материальных и человеческих потерь. Заблаговременное выявление угроз способствует сохранению материальных ценностей, уменьшению времени на их восстановление и, что важнее, спасению человеческих жизней. В связи с этим, предлагается новый подход к прогнозированию пожароопасных ситуаций: алгоритм обучения модели прогнозирования пожароопасных ситуаций, а также алгоритм прогнозирования пожароопасных ситуаций, которые разрабатываются на моделях машинного обучения таких как рекуррентные нейронные сети, случайный лес, оптимизационные деревья, авторегрессионные нейронные сети и т.д. В исследовании предлагается рассмотреть алгоритмы прогнозирования пожароопасных ситуаций, разработанных на основе анализа существующих алгоритмов прогнозирования, включая методы на основе машинного обучения, статистические модели и симуляционные подходы, учитывая их преимущества и недостатки, показатели точности. Результаты исследования разработанных алгоритмов показывают, что они способны с точностью 93.33% предсказывать значение температуры снаружи датчика на основе тестовых данных с комплекса взаимосвязанных пожарных датчиков, с ошибками MAE = 1.72, MSE = 2.95 в нештатном режиме на тестовых данных, и с точностью 92.85% температура внутри датчика, ошибки MAE = 1.66, MSE = 2.75. Точность на тестовых данных в штатном режиме для температуры снаружи составили 96.27%, ошибки MAE = 1.22, MSE = 1.48, а точность прогнозирования температуры внутри составила 96.16%, ошибки MAE = 1.24, MSE = 1.53. Для тестовой выборки в 500 000 отсчетов, ошибки спрогнозированной температуры снаружи составили: MAE = 1.82, а MSE = 3.31, а точность составили 91.78%. Ошибки спрогнозированной температуры внутри (temp2_inside) составили: MAE = 1.89, а MSE = 3.57, а точность составили 91.35%.

Пожарный датчик; Python; алгоритм прогнозирования; алгоритм обучения; прогнозирование; анализ алгоритмов прогнозирования.

S. Singh, A.V. Pribylskiy, E.Y. Kosenko

DEVELOPMENT AND RESEARCH OF ALGORITHMS FOR FORECASTING FIRE HAZARDOUS SITUATIONS

Early detection of fire hazard situations is a critical aspect of ensuring safety, as it helps to minimize the risk of material and human losses. Early detection of threats helps to preserve material assets, reduce the time for their restoration and, more importantly, save human lives. In this regard, a new approach to predicting fire hazard situations is proposed: an algorithm for training a model for predicting fire hazard situations, as well as an algorithm for predicting fire hazard situations, which are developed on machine learning models such as recurrent neural networks, random forest, optimization trees, autoregressive neural networks, etc. The study proposes to consider algorithms for predicting fire hazard situations developed on the basis of an analysis of existing forecasting algorithms, including methods based on machine learning, statistical models and simulation approaches, taking into account their advantages and disadvantages, accuracy indicators. The results of the study of the developed algorithms show that they are capable of predicting the outside temperature value of the sensor with an accuracy of 93.33% based on the test data from a complex of interconnected fire sensors, with errors of $MAE = 1.72$, $MSE = 2.95$ in the abnormal mode on the test data, and with an accuracy of 92.85% for the temperature inside the sensor, errors $MAE = 1.66$, $MSE = 2.75$. The accuracy on the test data in the normal mode for the outside temperature was 96.27%, errors $MAE = 1.22$, $MSE = 1.48$, and the accuracy of predicting the inside temperature was 96.16%, errors $MAE = 1.24$, $MSE = 1.53$. For the test sample of 500,000 readings, the errors of the predicted outside temperature were: $MAE = 1.82$, and $MSE = 3.31$, and the accuracy was 91.78%. The errors of the predicted temperature inside ($temp2_inside$) were: $MAE = 1.89$, and $MSE = 3.57$, and the accuracy was 91.35%.

Fire detector; Python; forecasting algorithm; learning algorithm; forecasting; forecasting algorithm analysis.

Введение. Прогнозирование пожароопасных ситуаций является одной из актуальных задач в области пожарной безопасности не имеющих единого и точного решения. Своевременное обнаружение и предупреждение пожаров позволяет значительно снизить риск человеческих жертв и материальных потерь. Традиционные системы мониторинга, основанные на пороговых значениях датчиков дыма и температуры, способны реагировать только на уже возникшие угрозы, кроме того быстродействие этих систем также зависит от целого ряда факторов: количество датчиков, расположения датчиков, чувствительность датчиков, алгоритма порогового обнаружения и т.д.. При этом даже системы с высокой чувствительностью и скоростью обнаружения не способны прогнозировать будущие угрозы, информировать и выполнять действия по предотвращению возникновения пожароопасной ситуации. В связи с этим, становятся актуальными разработка и исследование новых подходов к прогнозированию пожароопасных ситуаций как на основе классических статистических методов, так и набирающих популярность моделях машинного обучения, например использование рекуррентных нейронных сетей (RNN) или моделей на основе деревьев решений.

Рекуррентные нейронные сети обладают способностью моделировать временные зависимости и учитывать динамические изменения параметров, что делает их подходящими для анализа данных с пожарных датчиков. Деревья решений, в свою очередь, предоставляют интерпретируемую модель, которая может быть адаптирована и настроена для конкретных условий, что может быть полезно при эксплуатации датчиков в определенной области с изменяющимися параметрами, значение, которых можно приближенно рассчитать по составленной древовидной логической сети.

Однако прежде, чем приступать к выбору архитектуры, параметров, топологии модели машинного обучения, следует выбрать и разработать алгоритм обучения модели прогнозирования, таким образом, чтобы он подходил для большинства моделей машинного обучения применяемых для решения задачи прогнозирования с минимальными изменениями в структуре алгоритма. Следующим, но не менее важным этапом, является выбор самого алгоритма прогнозирования, потому как даже обладая точной и быстрой обученной моделью, при непроработанном алгоритме прогнозирования, можно по итогу

получить спрогнозированные данные с неудовлетворительной точностью. Алгоритм прогнозирования позволяет учитывать, параметры, факторы, которые сложно или невозможно учесть при обучении модели прогнозирования, например фильтрация выходных данных, сложная задача машинного обучения, в некоторых ситуациях – нерешаемая.

В данной статье рассматриваются исследования касающиеся разработки алгоритма прогнозирования и алгоритма обучения модели прогнозирования на основе рекуррентных нейронных сетей и деревьев решений, способных своевременно и с определенной точностью предсказывать пожароопасные ситуации на основе данных с комплекса взаимосвязанных пожарных датчиков.

Анализ аналогов. На сегодняшний день алгоритмы прогнозирования пожароопасных ситуаций преимущественно применяются для прогнозирования пожаров в лесах, а также для прогнозирования площади будущего расширения зоны пожара. Тем не менее, существующие решение в своей основе используют разные подходы к прогнозированию пожаров. В рамках текущего анализа условно можно выделить следующие подходы к построению алгоритмов прогнозирования:

- 1) на основе машинного обучения [1];
- 2) статистические модели [2];
- 3) на основе симуляции [3].

В работе [1] проведено исследование прогнозирования риска лесных пожаров в районе Хэйхэ (провинция Хэйлуцзян, Китай) с использованием алгоритмов машинного обучения Random Forest (RF) и Backpropagation Neural Network (BPNN). Анализируются ключевые факторы, влияющие на возникновение пожаров, такие как метеорологические данные, топография и человеческая активность. Модели оцениваются с помощью ROC-кривой.

Алгоритм на основе RF работает следующим образом: пусть обучающий набор данных $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. Финальное классификационное решение $H(x)$ принимается как:

$$H(x) = \arg \max_{y \in Y} \sum_{i=1}^{n_{tree}} I(h_i(x) = y), \quad (1)$$

где $h_i(x)$ – слабый классификатор, а $I(h_i(x) = y)$ – индикаторная функция, x_i, y_i – входные и выходные данные соответственно. Для заданных классификационных моделей $m_1(x), m_2(x), \dots, m_k(x)$ маргинальная функция вычисляется как:

$$mg(x, y) = a\vartheta_k I(m_k(x) = y) - \max_{j \neq k} a\vartheta_k I(m_k(x) = j), \quad (2)$$

где $a\vartheta_k$ – среднее значение индикаторной функции для k-го классификатора, j – альтернативный класс, отличный от k.

Ошибки определяются как:

$$PE = P_{x,y}(mg(x, y) < 0), \quad (3)$$

где $P_{x,y}$ – вероятность того, что маргинальная функция $mg(x, y)$ отрицательна, что указывает на ошибку классификации.

BPNN работает следующим образом: входные обучающие образцы проходят через сеть слой за слоем, на выходном слое сравнивается ошибка между фактическим и ожидаемым выходным значением, ошибка обратно распространяется, и веса и пороги корректируются слой за слоем, пока ошибка не уменьшится до заданного уровня точности. BPNN можно оптимизировать, изменяя топологию сети, скорость обучения, начальные веса и пороги.

К преимуществам метода можно отнести:

- ◆ высокая точность прогнозирования (оба метода показали area under the curve (AUC) > 0,93);
- ◆ способность работать с большими и сложными данными;
- ◆ устойчивость к шуму и выбросам;

К недостаткам метода можно отнести:

- ◆ необходимость большого объема данных для обучения;
- ◆ сложность интерпретации результатов (особенно для RF);
- ◆ риск переобучения при неправильной настройке (особенно для BPNN).

Исследование [2] посвящено оценке двух подходов к прогнозированию экстремальной опасности пожаров в Австралии: динамической модели климатической симуляции Australian Community Climate and Earth-System Simulator (ACCESS-S2) и статистической модели, основанной на климатических драйверах (фактор климата). Анализируются три значимых случая экстремальных пожаров: Canberra (2003), Black Saturday (2009) и Pinery (2015). Статистическая модель, построенная с использованием линейной логистической регрессии, показала более высокую вероятность обнаружения (POD) по сравнению с динамической моделью, но также имела более высокий уровень ложных срабатываний (FAR).

Статистическая модель основана на линейной логистической регрессии, которая использует климатические индексы для прогнозирования вероятности экстремальных значений индекса поведения огня (FBI). Модель строится на основе данных о климатических драйверах, таких как MJO, ENSO, IOD, SAM, блокирование потока и субтропический хребет. Вероятность экстремальных пожаров определяется как бинарная переменная (да/нет) на основе порога 0,1.

К преимуществам метода можно отнести:

- ◆ простота реализации модели, а так интерпретируемость;
- ◆ статистическая модель показала вероятность обнаружения (POD) в диапазоне 0,57–0,87, что значительно выше, чем у динамической модели (0,03–0,57);
- ◆ возможность использования для долгосрочных прогнозов (до нескольких месяцев).

К недостаткам метода можно отнести:

- ◆ высокий уровень ложных срабатываний (FAR), в диапазоне 0,44–0,66;
- ◆ ограниченная способность учитывать нелинейные зависимости между климатическими драйверами и пожарами;
- ◆ зависимость от точности данных о климатических драйверах.

Проанализируем статью [3]. Алгоритм прогнозирования распространения лесных пожаров, использованный в исследовании, основан на модели WRF-SFIRE. Эта модель объединяет метеорологическое моделирование с моделированием распространения пожаров, что позволяет учитывать взаимодействие между атмосферными условиями и динамикой пожара. Основные этапы алгоритма:

1) Сбор данных:

- ◆ метеорологические данные: используются прогнозируемые данные от Global Forecast System (GFS) с разрешением 0.25 градуса. Эти данные включают температуру, скорость и направление ветра, относительную влажность и другие метеорологические параметры;
- ◆ топливные параметры: собираются данные о глубине топлива, нагрузке топлива и содержании влаги на основе полевых образцов, взятых из 23 участков. Топливо классифицируется по размерам (1h, 10h, 100h, 1000h) и анализируется для определения его свойств, таких как содержание влаги и соотношение поверхности к объему (SAVR);
- ◆ данные о рельефе: используется цифровая модель рельефа (DEM) с разрешением 12.5 м, полученная от спутниковых данных ALOS PALSAR;
- ◆ карта землепользования и растительности (LULC): используется карта в масштабе 1:10,000, предоставленная Департаментом науки и технологий Сиккима, Индия.

2) Подготовка данных:

- ◆ метеорологические данные динамически уточняются с помощью вложенных доменов в модели WRF, что позволяет достичь высокого разрешения (до 40 м) для моделирования распространения пожара;
- ◆ топливные параметры связываются с классами растительности на карте LULC, что позволяет создать карту топлива с разрешением 40 м.

3) Моделирование распространения пожара:

- ◆ модель Ротермеля: используется для оценки скорости распространения пожара. Эта модель основана на балансе тепла и учитывает параметры топлива (глубина, нагрузка, содержание влаги) и метеорологические условия (скорость ветра, уклон местности);
- ◆ метод уровней множеств: применяется для моделирования распространения пожара в сетке с высоким разрешением. Этот метод позволяет отслеживать эволюцию пожара во времени и пространстве с использованием уравнений в частных производных.

4) Инициация пожара – в качестве точки инициации пожара используются данные о текущих пожарах, полученные от спутниковых сенсоров VIIRS. Эти данные синхронизируются с прогнозируемыми метеорологическими данными для динамического моделирования распространения пожара.

5) Прогнозирование – модель выполняет 24-часовой прогноз распространения пожара, учитывая взаимодействие между атмосферными условиями и пожаром. Результаты прогноза включают карты выгоревшей территории с разрешением 40 м.

6) Валидация – прогнозируемая выгоревшая территория сравнивается с фактической выгоревшей территорией, определенной на основе спутниковых данных Sentinel 2. Это позволяет оценить точность модели.

К преимуществам метода можно отнести:

- ◆ модель учитывает динамическое взаимодействие между атмосферными условиями и поведением пожара, что позволяет более точно прогнозировать изменения в направлении и скорости распространения огня;
- ◆ точность совпадения прогнозируемой и фактической выгоревшей территории составили 0.85–0.9, за исключением небольших расхождений в юго-западном направлении.

К недостаткам метода можно отнести:

- ◆ для выполнения 24-часового прогноза потребовалось 2 часа 11 минут на суперкомпьютере PARAM SEVA с использованием 240 ядер, что указывает на значительные вычислительные ресурсы;
- ◆ в некоторых регионах может отсутствовать достаточное количество данных о топливе, что снижает точность прогноза. Например, в исследовании использовалась модифицированная модель Андерсона из-за недостатка локальных данных;
- ◆ в горных районах оползни могут изменять ландшафт и действовать как барьеры для пожаров, но они не всегда учитываются в картах растительности, что снижает точность прогноза.

Приведем результаты анализа в виде табл. 1.

Таблица 1

Сравнительная анализ алгоритмов прогнозирования

Алгоритм	Точность	Работа в режиме реального времени	Устойчивость к шумам и выбросам	Нахождение сложных нелинейных зависимостей
На основе машинного обучения	Высокая 0.93 и выше	да	да	да
Статистические модели	Средняя (0.57-0.87)	да	нет	нет
На основе симуляции	Высокая 0.85-0.9	Да (но, самая высокая ресурсозатратность)	да	да

Анализ исследований [1–3] показывает, что машинное обучение может обеспечивать высокую точность прогнозирования, но чем, сложнее объект прогнозирования и задачи прогнозирования, тем более мощные вычислительные ресурсы могут потребоваться для обуче-

ния, совместно с большими объема обучающих и тестирующих данных. Статистические модели, как правило используются для анализа и предсказания будущих пожаров с большой длиной предсказаний, при этом они не применимы для работы в режиме реального времени, в виду своей низкой точностью и невозможностью к адаптации к изменяющимся данным. Симуляционные модели позволяют с помощью симуляции прогнозировать развитие площади пожаров, визуализируя его на понятном даже обычному пользователю изображении, однако такие системы требуют построение детальной и адекватной математической модели объекта управления, что в некоторых ситуациях является невозможным, кроме того для моделирования пожара требуется гораздо больше вычислительных ресурсов, чем для прогнозирования пожаров с помощью машинного обучения или статистическим методом. Для универсального пожарного датчика [4–6] требуется подход, позволяющий прогнозировать пожароопасные ситуации в режиме реального времени с минимальным количеством запросов на настройку параметров алгоритмов, среди рассмотренных алгоритмов этим требованиям соответствуют алгоритмы прогнозирования на основе машинного обучения, которая позволяет находить скрытые функциональные зависимости в данных, тем самым прогнозирование данных основывается не только на имеющихся статистических данных, но и с учетом найденных скрытых зависимостей, которые интерпретируются алгоритмами машинного обучения, как функционально-весовыми (взвешенные зависимости), логически-весовыми (взвешенные логические условия) связями между параметрами.

Описание универсального датчика в виде модели вход-выход. Универсальный пожарный датчик представляет из себя комплекс взаимосвязанных датчиков [7, 8], с настраиваемыми параметрами, который в зависимости от настроек, входных потоков, возмущающих воздействий генерирует комплексный выходной сигнал. Параметрическая модель универсального датчика представлена на рис. 1.



Рис. 1. Параметрическая модель универсального пожарного датчика

Входящие преобразуемые информационные потоки:

- ◆ температура – измеряется из окружающей среды в месте установки, для этой цели используются 8 датчиков температура разной инерционности, из которых 1 с самой низкой инерционностью и 1 с самой высокой инерционностью (0–80 градусов).
- ◆ задымленность по ультрафиолетовому спектра – обнаруживается и измеряется с помощью трех датчиков задымленности ультрафиолетового спектра (0–65 556 аналоговых отсчетов);
- ◆ задымленность по инфракрасному спектра – обнаруживается и измеряется с помощью трех датчиков задымленности инфракрасного спектра (0–65 556 аналоговых отсчетов).

Возмущающие воздействия:

- ◆ электромагнитные помехи – внешние помехи, которые могут приводить к искажению [9] показанию датчика, интенсивность помех: 0–100 мВ (милливольт) для слабых помех, 100–1000 мВ для умеренных помех, >1 В (вольт) для сильных помех, частота помех: 50–60 Гц (частота сетевого напряжения), 100–500 МГц и выше (радиочастоты), 2.4–5 ГГц (Wi-Fi, Bluetooth);

- ◆ вибрации – возникают при размещении в агрессивных средах, например в вагоне метрополитена постоянно на датчик воздействует вибрация [10], амплитуда вибраций: 0–5 м/с² (миллиметры на секунду в квадрате) для слабых вибраций, 5–20 м/с² для умеренных вибраций, >20 м/с² для сильных вибраций, частота вибраций: 1–10 Гц (низкочастотные вибрации), 10–100 Гц (среднечастотные вибрации), >100 Гц (высокочастотные вибрации);

- ◆ влажность – искажают показания датчиков задымленности, в метрополитенах, например, это может быть туман, относительная влажность: 0–30% для сухого воздуха, 30–70% для нормальных условий, 70–100% для высокой влажности, абсолютная влажность: 0–10 г/м³ (граммов воды на кубический метр) для сухого воздуха, 10–20 г/м³ для нормальных условий, >20 г/м³ для высокой влажности.

Управляющие входы, параметры датчиков такие как: чувствительность (настраиваемый параметр для датчиков, 0-2 градуса для температурных датчиков, 0.1 -1% отсчетов для датчиков задымленности) [11, 12], пороги срабатывания (настраивается программно в пределах диапазона измерений), частота опроса датчика (настраивается, по умолчанию 5 раз в секунду), частота считывания информации датчиков (настраивается, по умолчанию 5 раз в секунду).

Выходы:

- ◆ преобразованная информация – обработанные данные с датчиком, передающиеся на центр обработки информации (контроллер, компьютер, диспетчерская и т.д.), в котором в зависимости от предусмотренных алгоритмов выполняются определенные предписания (Спрогнозированные значения, температур, задымленности, а также в некоторых случаях статистические параметры: дисперсия, производная определенного параметра, математическое ожидание и т.д.)

- ◆ средство пожаротушения – сигналы для активации средств пожаротушения, таких как спринклерные системы, огнетушители и т.д. (сигнал активации 0 или 1);

- ◆ индикаторы состояния – визуальные и звуковые сигналы, информирующие о текущем состоянии датчика (сигнал активации 0 или 1).

Алгоритм обучение и тестирования моделей прогнозирования пожароопасных ситуаций. Одним из основных и важных составляющих решения задачи прогнозирования данных, является выбор и описание алгоритма прогнозирования, которые определяют набор инструкций, которые необходимы для достижения поставленной цели, при этом алгоритм должен предусматривать возможность использования различных моделей совместно с представленным набором действий. В рамках статьи ниже предложен алгоритм прогнозирования и алгоритм обучения моделей прогнозирования, которые в качестве моделей использует RNN (recurrent neural network) и их модификации, а также модели на основе деревьев решений. Алгоритм обучения модели прогнозирования представлен на рис. 2,а.

На рис. 2,а используются 4 функции «Стандартизация формата времени», «Добавление новых параметров», «Разделение данных на тестовые и обучающие», «Выбор признаков для обучения», алгоритм, которых представлен ниже.

Задача алгоритма «Стандартизация формата времени» (рис. 2,б) заключается в стандартизации формата времени [13, 14] считанных данных, т.е. приведение к общему формату. Целевой формат имеет вид (4):

$$data_{time} == DD.MM.YYYY HH:MM:SS, \quad (4)$$

где DD – день месяца (01-31), MM – месяц (01-12), YYYY – год (четырёхзначное число), HH – часы в 24-часовом формате (00-23), MM – минуты (00-59), SS – секунды (00-59).

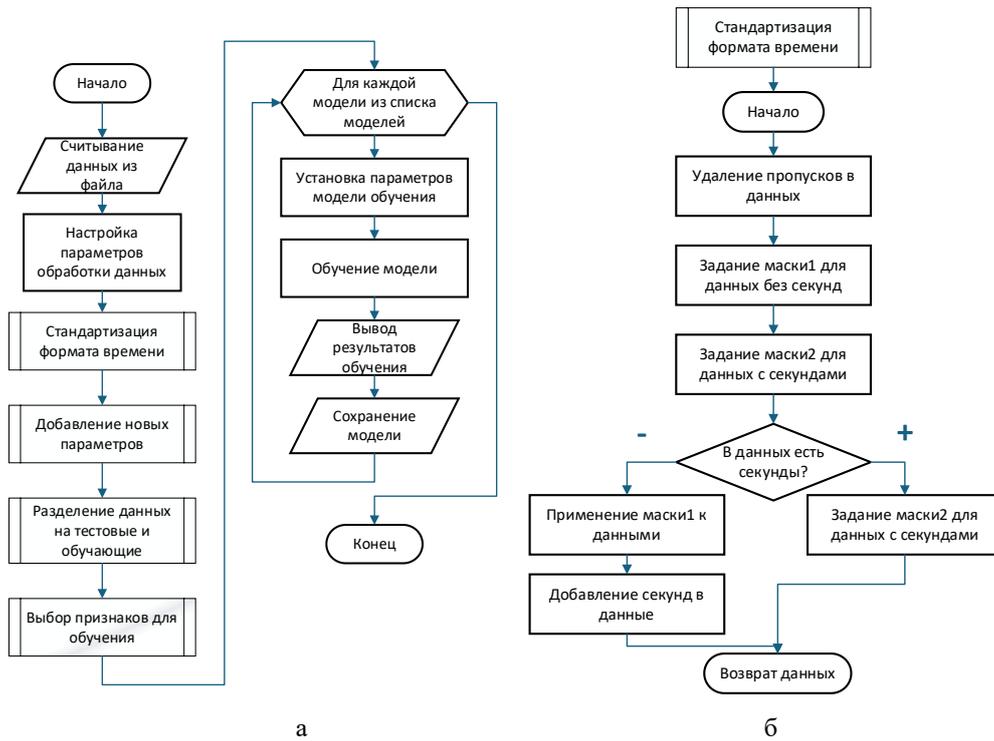


Рис. 2. а – алгоритм обучения модели прогнозирования, б – алгоритм функции «Стандартизация формата времени»

Алгоритм добавление новых параметров представлен на рис. 3.

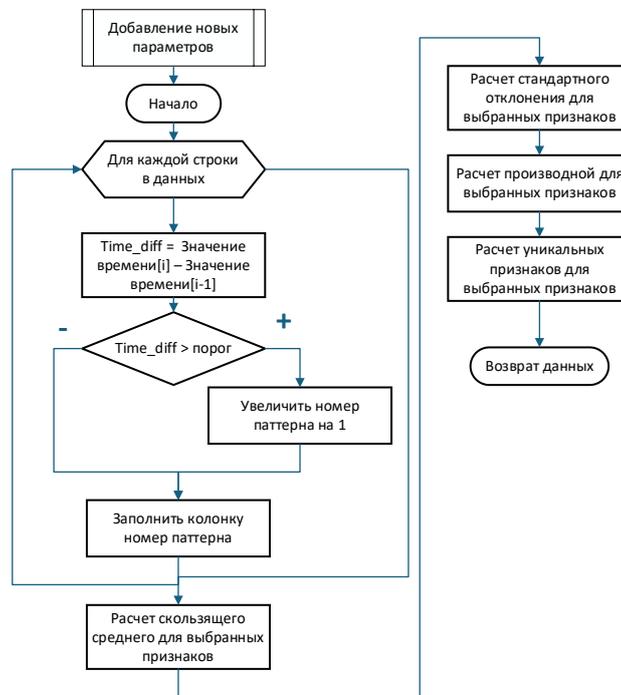


Рис. 3. Алгоритм функции «Добавление новых параметров»

Функция «Добавление новых параметров» добавляет признаки в данные, которые необходимы для обучения модели прогнозирования конкретного параметра, при этом, добавление дополнительных признаков может повысить точность модели.

Задача функции «Преобразование в формат для обучения» (рис. 4) заключается в преобразовании данных в «supervised data», которая будет хранить в себе историю о прошлых показаниях для каждой строки, тем самым добавляя «память о прошлом» в базу данных.

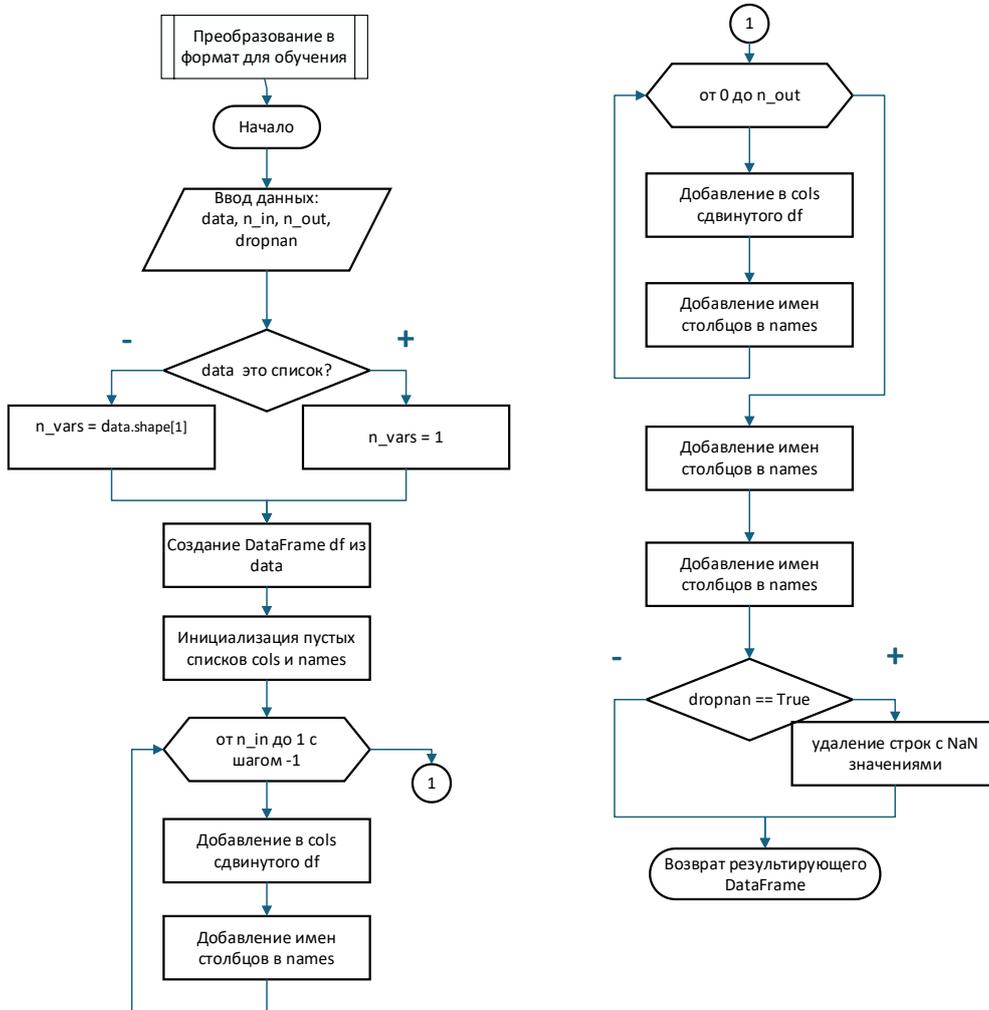


Рис. 4. Алгоритм функции «Преобразование в формат для обучения»

Следующим этапом обучение модели прогнозирования, является применением обученной модели на тестовых данных. Перед этим необходимо вызвать функцию разделения данных (рис. 5,а), функцию выбора признаков обучения (рис. 5,б). Алгоритм прогнозирования представлен на рис. 5,в.

Функция «Разделение данных на тестовые и обучающие», разделяет данные на тренировочные – используемые для обучения модели и тестовые – используемые для проверки работоспособности модели.

Функция «Выбор признаков для обучения» преобразовывает данные Test и Train, оставляя в ней только признаки необходимые для обучения модели.

Раскроем подробнее алгоритм функций «Прогнозировать будущие значения» (рис. 6).

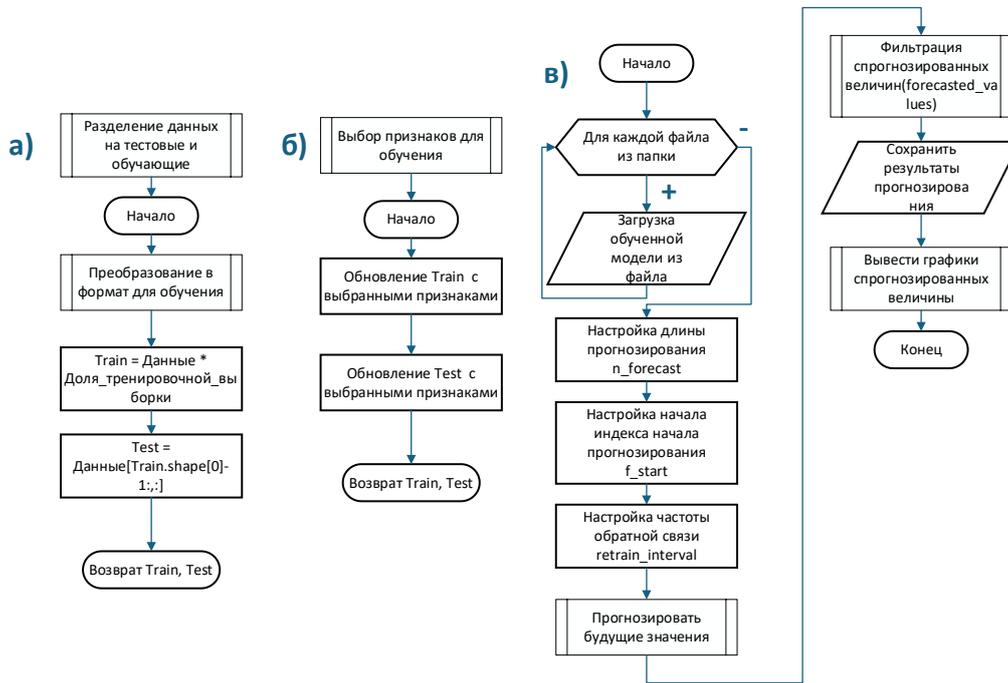


Рис. 5. Алгоритм функций: а – функция разделения данных, б – функция выбора признака для обучения, в – алгоритм прогнозирования

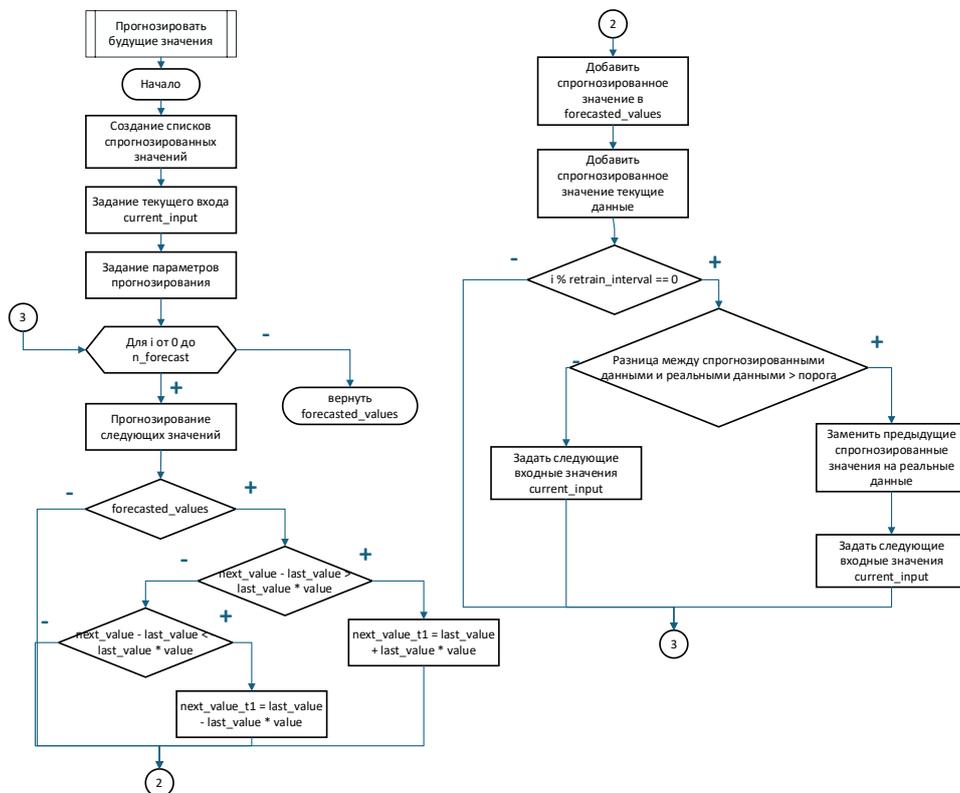


Рис. 6. Алгоритм функции «Прогнозировать будущие значения»

На рис. 6 представлена часть алгоритма функции «Прогнозировать будущие значения», в ней описана инициализация переменных, задание входных значений, а также приведено условие коррекции спрогнозированных значений.

После блока ссылка (под №2) следующая идущая часть алгоритма выполняет задачу обратной связи спрогнозированных данных с реальными данным, а именно корректирует прошлые спрогнозированные значения заменяя их на реальные данные, если разница между ними превышает определенное значение.

Раскроем алгоритм функции «Фильтрация спрогнозированных величин» (рис. 7)

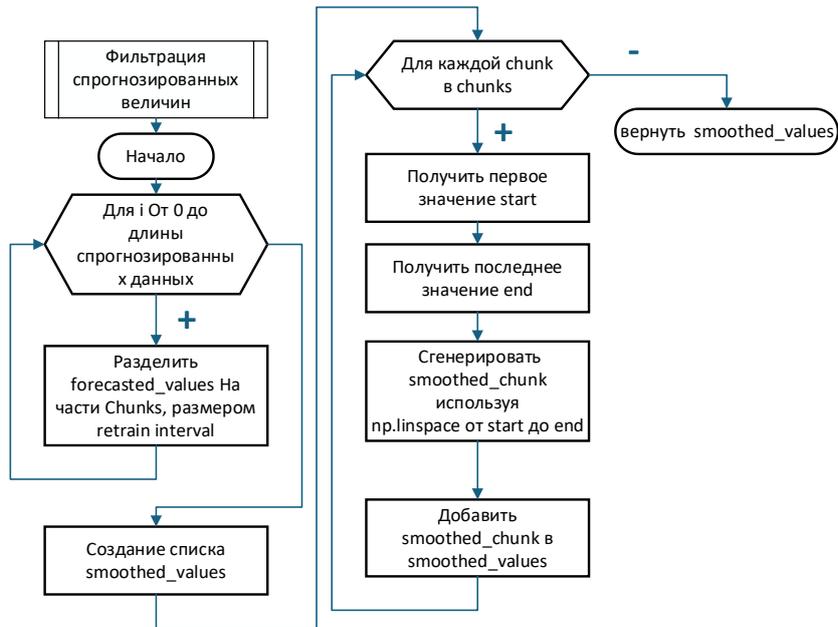


Рис. 7. Алгоритм функции «Фильтрация спрогнозированных величин»

Задача функции «Фильтрация спрогнозированных величин» (рис. 7) заключается в сглаживании спрогнозированных значений для упрощения дальнейшего анализа и обработки данных, например для классификации данных, с целью выявления пожароопасных ситуаций.

Подготовка данных. В качестве данных для обучения и тестирования алгоритма прогнозирования, используются реальные данные выгруженный из базы данных метрополитенов, в которых установлены универсальные пожарные датчики. Перед использование этих данных в представленном алгоритме необходимо выполнить парсинг данных для того, чтобы его можно было использовать в моделях прогнозирования. После парсинга данные имеют вид, представленный на рис. 8.

	A	B	C	D	E	F	G	H	I
1	date_time	temp1_outside	temp2_inside	Smoke_red_1	Smoke_red_2	Smoke_red_3	Smoke_blue_1	Smoke_blue_2	Smoke_blue_3
2	29.01.2021 0:00:00	16.61	16	60121	60124	60118	60117	60121	60120
3	29.01.2021 0:00:05	16.7	15.97	60120	60121	60120	60117	60122	60119
4	29.01.2021 0:00:10	16.66	15.98	60120	60119	60123	60117	60117	60120
5	29.01.2021 0:00:15	16.63	15.96	60119	60123	60122	60119	60117	60117
6	29.01.2021 0:00:20	16.6	15.98	60122	60123	60120	60119	60120	60118
7	29.01.2021 0:00:25	16.56	15.96	60122	60119	60118	60117	60124	60122
8	29.01.2021 0:00:30	16.57	16	60121	60122	60124	60124	60117	60122
9	29.01.2021 0:00:35	16.56	15.99	60121	60119	60124	60118	60123	60122
10	29.01.2021 0:00:40	16.58	15.96	60120	60120	60122	60121	60119	60124
11	29.01.2021 0:00:45	16.55	15.98	60119	60124	60122	60119	60119	60121
12	29.01.2021 0:00:50	16.56	15.98	60121	60121	60120	60119	60124	60119
13	29.01.2021 0:00:55	16.52	15.98	60124	60121	60121	60118	60118	60120
14	29.01.2021 0:01:00	16.51	15.98	60119	60119	60123	60123	60117	60121

Рис. 8. Пример обработанных данных выгруженных из базы данных метрополитена

В зависимости от используемой модели данных могут вводиться новые параметры, полученные как путем математических вычислений с текущими данными (среднее значение, дисперсия, производная и т.д.), так и параметры на основе экспертных знаний или результатов модели прогнозирования (история показаний, паттерн данных и т.д.).

Тестирование алгоритмов. Тестирование представленных алгоритмов осуществляется с применением модели прогнозирования на основе дерева решения Random Forest (Случайные лес) [15–18], в качестве прогнозируемых параметров выбраны temp1_outside и temp2_inside, обучающая выборка составила 1 200 000 отсчетов В результате обучение получены модели, представленные в табл. 2.

Таблица 2

Результат обучение моделей

Название модели	Объем занимаемой памяти	Параметр	Диапазоны значений
model_rf_temp1.pkl	404 МБ	Температура снаружи датчика	0 – 80 °С
model_rf_temp2.pkl	411 МБ	Температура внутри датчика	0 – 80 °С
model_rf_blue1.pkl	69 МБ	Задымленность по ультрафиолетовому каналу	0 – 65 556
model_rf_red.pkl	71 МБ	Задымленность по инфракрасному каналу	0 – 65 556

Применим полученные нормализованные данные в ранее представленных алгоритмах прогнозирования (рис. 5–7).

В качестве метрики оценки точности прогнозирования будет использоваться средняя абсолютная ошибка и среднеквадратичная ошибка [19, 20].

Формула средней абсолютной ошибки MAE (5):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (5)$$

где y_i – фактическое значение, \hat{y}_i – прогнозируемое значение, n – количество наблюдений.

Среднеквадратичная ошибка MSE (6):

$$MSE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|^2, \quad (6)$$

Для тестовой выборки в 500 000 отсчетов, ошибки спрогнозированной температуры снаружи (temp1_outside) составили: MAE = 1.82, а MSE = 3.31, а точность составили 91.78%, что для задачи прогнозирования температуры является нормальным показателем [21–23]. Ошибки спрогнозированной температуры внутри (temp2_inside) составили: MAE = 1.89, а MSE = 3.57, а точность составили 91.35%, что для задачи прогнозирования температуры является нормальным показателем.

Критическим значением будем считать MAE >4, MSE >8, точность ниже 85%.

Рассмотрим прогнозирования значения в двух ситуациях, в нормальном режиме (штатный режим работы) – пожароопасные ситуации отсутствуют и в ситуации возникновения пожароопасной ситуации (нештатный ситуация), для того чтобы можно было оценить, результаты работы комплексного алгоритма прогнозирования в различных ситуациях. Сначала рассмотрим штатный режим работу (рис. 9, 10)

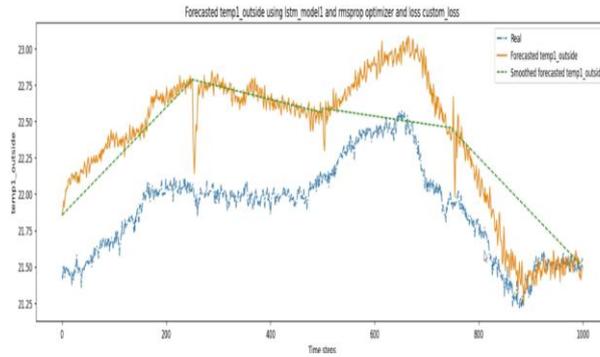


Рис. 9. График прогнозирование температуры $temp1_outside$ в штатном режиме работы

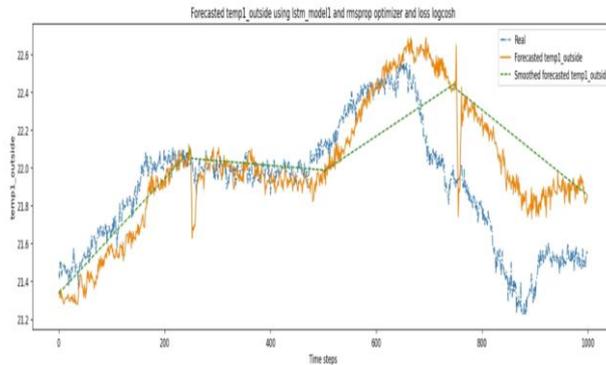


Рис. 10. График прогнозирование температуры $temp2_inside$ в штатном режиме работы

Анализируя графики полученный на рис. 9, 10, можно сделать вывод, что прогнозируемые значения соответствуют реальным данным с ошибками $MAE = 1.22$, $MSE = 1.48$ для рис. 9 и $MAE = 1.24$, $MSE = 1.53$ для рис. 10, точность прогнозирования составила 96.27% и 96.16% соответственно. Стоит также отметить, что отфильтрованные спрогнозированные значение представляют из себя аппроксимированную линейную функцию спрогнозированных значений, что отражается на локальных значениях температуры. Чтобы повысить локальную точность отфильтрованных спрогнозированных значений температуры, следует применить другой метод фильтрации данных.

Рассмотрим прогнозирование данных на участке с нештатной ситуацией (рис. 11, 12).

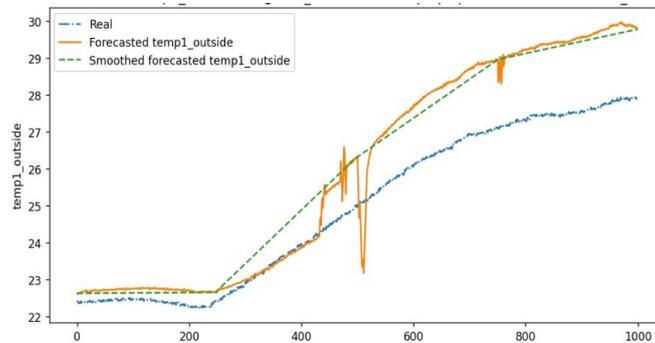


Рис. 11. График прогнозирование температуры $temp1_outside$ в штатном режиме работы

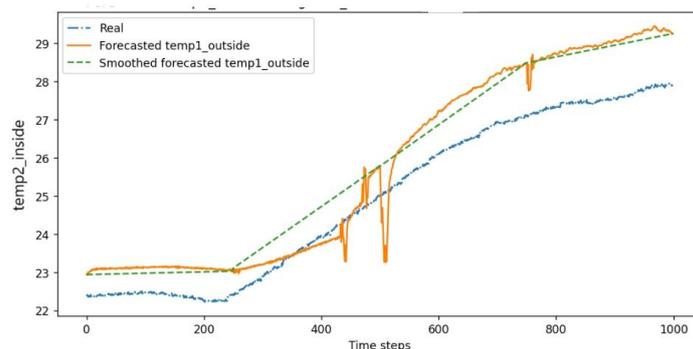


Рис. 12. График прогнозирования температуры $temp2_inside$ в штатном режиме работы

Анализируя графики полученные на рис. 11, 12, можно сделать вывод, что прогнозируемые значения соответствуют реальным данным с ошибками $MAE = 1.72$, $MSE = 2.95$ для рис. 11 и $MAE = 1.66$, $MSE = 2.75$ для рис. 12, точность прогнозирования составили 93.33% и 92.85% соответственно.

Очевидно, что при возникновении нештатной ситуации ошибки MSE и MAE больше, чем в штатном режиме, однако они не выходят за установленные выше пороги.

Также можно протестировать совместную работу классификатора [5, 6] пожароопасных ситуаций в совместной работе с разработанным и реализованным алгоритмом прогнозирования (рис. 13).

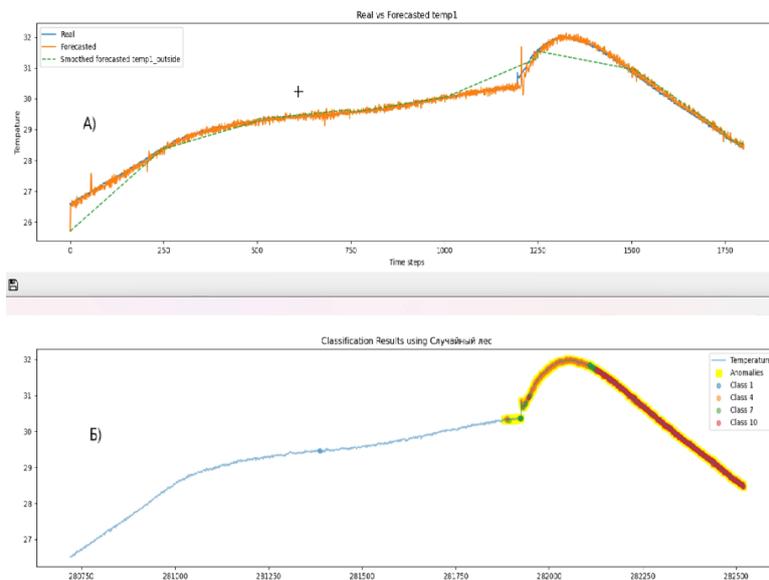


Рис. 13. А – график прогнозирования пожароопасных ситуаций, Б – график классификация пожароопасных ситуаций на основании спрогнозированных данных

На рис. 13,А, Real – график изменения реальной температуры, Forecasted спрогнозированной значение температуры, Smoothed forecasted temp1_outside – отфильтрованное спрогнозированное значение температуры. На рис. 13,Б, Class 1-n, класс распознанной пожароопасной ситуации, Anomalies – реальные пожароопасные ситуаций, Temperature – температура.

Обсуждение. В данной статье предложены алгоритмы прогнозирования и обучения моделей прогнозирования пожароопасных ситуаций.

Результаты тестирования предложенных алгоритмов с применением модели Random forest на реальных данных показали, что обученные по предложенному алгоритму (рис. 2–5) модели способны используя алгоритм (рис. 5–7) прогнозировать пожароопасные ситуации со следующими показателями точности, для тестовой выборки в 500,000 отсчетов, при прогнозировании температуры снаружи, средняя абсолютная ошибка (MAE) составила 1.82, а среднеквадратичная ошибка (MSE) – 3.31, точность прогнозирования 91.78%. Ошибки спрогнозированной температуры внутри (temp2_inside) составили: MAE = 1.89, а MSE = 3.57, а точность составили 91.35%, что для задачи прогнозирования температуры является нормальным показателем. Эти значения свидетельствуют о том, что модели не превышают критические значения показателей точности, приведенные выше.

Анализ графиков прогнозирования температуры в штатном и нештатном режимах работы показал, что модели способны адекватно реагировать на изменения условий. На примере температуры снаружи, в штатном режиме ошибки прогнозирования были минимальны (MAE = 1.22, MSE = 1.48), точность 96.26%, тогда как в нештатном режиме они несколько увеличивались (MAE = 1.72, MSE = 2.95), точность 93.33%, но тем не менее, показатели оставались в пределах допустимых значений.

Заключение. Предложенные алгоритмы обучения и прогнозирования демонстрируют точность не установленных выше критических значений, при их реализации на модели Random Forest, модель обрабатывает штатный и нештатный ситуации, не превышая заданные MAE и MSE, для тестовой выборки в 500 000 отсчетов, MAE составило 1.82, MSE - 3.31.

В дальнейшем планируется более глубокую оценку точности прогнозирования с использованием различных моделей и методов фильтрации данных, чтобы выбрать опорную модель и её параметры для дальнейшей реализации системы прогнозирования и классификации пожароопасных ситуаций.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Chao Gao, Honglei Lin, Haiqing Hu.* Forest-Fire-Risk Prediction Based on Random Forest and Backpropagation Neural Network of Heihe Area in Heilongjiang Province // *Forests.* – 2023. – DOI: 10.3390/f14020170.
2. *Marshall A.G., Crimp S., Cary G.J., & Harris S.* A Statistical Forecasting Model for Extremes of the Fire Behaviour Index in Australia // *Atmosphere.* – 2024. – <https://doi.org/10.3390/atmos15040470>.
3. *Sultan Md.A., Limboo N., Mukherjee A., Kharkar N., Islam S., Pokale S., Talekar S., & Khare M.* Operational Forest-Fire Spread Forecasting Using the WRF-SFIRE Model // *Remote Sensing.* – 2024. – 16 (13). – 2480. – <https://doi.org/10.3390/rs16132480>.
4. *Сингх С., Прибыльский А.В.* Синтез системы сверхбыстрого обнаружения пожароопасных ситуаций на основе комплекса взаимосвязанных датчиков // *Известие ЮФУ. Технические науки.* – 2024. – № 2. – С. 121-132.
5. *Сингх С., Прибыльский А.В.* Алгоритм классификации пожароопасных ситуаций на основе нейросетевых технологий // *Известия ЮФУ. Технические науки.* – 2024. – № 3. – С. 138-147.
6. *Сингх С., Прибыльский А.В.* Классификации пожароопасных ситуаций на основе сети Колмогорова-Арнольда // *Известие ЮФУ. Технические науки.* – 2024. – № 6. – С. 6-15.
7. *Саутин И.Г.* Противопожарная защита: технологии и решения // *Транспорт. Противопожарная защита. Пожарная автоматика. Средства спасения.* – 2018.
8. *Саутин И.Г.* Особое мнение. Можно ли доверить свою жизнь дымовому пожарному извещателю? // *Алгоритм безопасности.* – 2019. – № 6.
9. *Shchemelev V. & Ezhov Yu & Zub I.* The influence of external factors determining the operation of the positioning sensor in the condition of the far North // *AIP Conference Proceedings.* – 2023. – 2700. – 060015. 10.1063/5.0125043.
10. *El Abdi R., Labbé J., Le Strat F., & Carvou E.* Effect of Vibration Frequency on Mechanical Behavior of Automotive Sensor. – Springer, Cham, 2018. – P. 1-7. – https://doi.org/10.1007/978-3-319-96358-7_1.
11. NFPA 72: National Fire Alarm and Signaling Code. – 2019.
12. UL 268: Standard for Smoke Detectors for Fire Protective Signaling Systems. – 2019.

13. Khan Zanis Ali & Shin Donghwan & Bianculli Domenico & Briand Lionel. Impact of log parsing on deep learning-based anomaly detection // *Empirical Software Engineering*. – 2024. – 29. – 10.1007/s10664-024-10533-w.
14. Sellberg F & Buthke J & Sonne-Frederiksen Povl Filip & Nørkjær Gade Peter. Evaluating Four Types of Data Parsing Methods for Machine Learning Integration from Building Information Models. – 2022.
15. Fan Gaohun. Random Forest Algorithm for Forest Fire Prediction. – 2023. – 10.1007/978-981-99-4554-2_15.
16. Pande Chaitanya & Radwan Neyara & Heddham Salim & Othman Kaywan & Alshehri Fahad & Pal Subodh & Pramanik Malay. Forecasting of monthly air quality index and understanding the air pollution in the urban city, India based on machine learning models and cross-validation // *Journal of Atmospheric Chemistry*. – 2024. – 82. – P. 1-26. – 10.1007/s10874-024-09466-x.
17. Salman Hasan & Kalakech Ali & Steiti Amani. Random Forest Algorithm Overview // *Babylonian Journal of Machine Learning*. – 2024. – P. 69-79. – 10.58496/BJML/2024/007.
18. Bijan Ahmed & Al-Rahim Ali. Random Forest and Decision Tree Facies Classification Models for Well Log Data of the Mishrif Formation from Basrah Oil Company, Southern Iraq // *Iraqi Geological Journal*. – 2025. – 57. – P. 14-32. – 10.46717/igj.57.2E.2ms-2024-11-11.
19. Al-Janabia Samaher & AlShourbaji Ibrahim & Patel Ahmed. Applied Predicative Modeling to Improve Recommendation Systems for Forecasting of Fire Occurrences. – 2015.
20. Ramadhan Rafiq & Ashari Wahid. Performance Comparison of Random Forest and Decision Tree Algorithms for Anomaly Detection in Networks // *Journal of Applied Informatics and Computing*. – 2024. – 8. – P. 367-375. – 10.30871/jaic.v8i2.8492.
21. Li Liping. Comparative Research on Diabetes Influencing Factors Based on Random Forest and Decision Tree Models // *Highlights in Science, Engineering and Technology*. – 2023. – 72. – P. 231-242. – 10.54097/7m4x7j04.
22. Hu Yaowen. Comparison and Analysis of the Effectiveness of Linear Regression, Decision Tree, and Random Forest Models for Health Insurance Premium Forecasting // *Advances in Economics, Management and Political Sciences*. – 2024. – 79. – P. 347-353. – 10.54254/2754-1169/79/20241754.
23. Cansler C. & Wright Micah & Mantgem Phillip & Shearman Timothy & Varner J. & Hood Sharon. Drought before fire increases tree mortality after fire // *Ecosphere*. – 2024. – 15. – 10.1002/ecs2.70083.

REFERENCES

1. Chao Gao, Honglei Lin, Haiqing Hu. Forest-Fire-Risk Prediction Based on Random Forest and Backpropagation Neural Network of Heihe Area in Heilongjiang Province, *Forests*, 2023. DOI: 10.3390/f14020170.
2. Marshall A.G., Crimp S., Cary G.J., & Harris S. A Statistical Forecasting Model for Extremes of the Fire Behaviour Index in Australia, *Atmosphere*, 2024. Available at: <https://doi.org/10.3390/atmos15040470>.
3. Sultan Md.A., Limboo N., Mukherjee A., Kharkar N., Islam S., Pokale S., Talekar S., & Khare M. Operational Forest-Fire Spread Forecasting Using the WRF-SFIRE Model, *Remote Sensing*, 2024, 16 (13), 2480. Available at: <https://doi.org/10.3390/rs16132480>.
4. Singkh S., Pribyl'skiy A.V. Sintez sistemy sverkhbystrogo obnaruzheniya pozharoopasnykh situatsiy na osnove kompleksa vzaimosvyazannykh datchikov [Synthesis of a system for ultra-fast detection of fire hazardous situations based on a complex of interconnected sensors], *Izvestie YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2024, No. 2, pp. 121-132.
5. Singkh S., Pribyl'skiy A.V. Algoritm klassifikatsii pozharoopasnykh situatsiy na osnove neyrosetevykh tekhnologiy [Algorithm for classification of fire hazardous situations based on neural network technologies], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2024, No. 3, pp. 138-147.
6. Singkh S., Pribyl'skiy A.V. Klassifikatsii pozharoopasnykh situatsiy na osnove seti Kolmogorova-Arnol'da [Classifications of fire hazardous situations based on the Kolmogorov-Arnold network], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2024, No. 6, pp. 6-15.
7. Sautin I.G. Protivopozharnaya zashchita: tekhnologii i resheniya [Fire protection: technologies and solutions], *Transport. Protivopozharnaya zashchita. Pozharnaya avtomatika. Sredstva spaseniya* [Transport. Fire protection. Fire automation. Rescue equipment], 2018.
8. Sautin I.G. Osoboe mnenie. Mozhno li doverit' svoyu zhizn' dymovomu pozharnomu izveshchatel'yu? [Dissenting opinion. Can you trust your life to a smoke alarm?], *Algoritm bezopasnosti* [Security Algorithm], 2019, No. 6.
9. Shchemelev V. & Ezhov Yu & Zub I. The influence of external factors determining the operation of the positioning sensor in the condition of the far North, *AIP Conference Proceedings*, 2023, 2700. 060015. 10.1063/5.0125043.
10. El Abdi R., Labbé J., Le Strat F., & Carvou E. Effect of Vibration Frequency on Mechanical Behavior of Automotive Sensor. Springer, Cham, 2018, pp. 1-7. Available at: https://doi.org/10.1007/978-3-319-96358-7_1.
11. NFPA 72: National Fire Alarm and Signaling Code, 2019.

12. UL 268: Standard for Smoke Detectors for Fire Protective Signaling Systems, 2019.
13. Khan Zanis Ali & Shin Donghwan & Bianculli Domenico & Briand Lionel. Impact of log parsing on deep learning-based anomaly detection, *Empirical Software Engineering*, 2024, 29. 10.1007/s10664-024-10533-w.
14. Sellberg F & Buthke J & Sonne-Frederiksen Povl Filip & Nørkjær Gade Peter. Evaluating Four Types of Data Parsing Methods for Machine Learning Integration from Building Information Models, 2022.
15. Fan Gaolun. Random Forest Algorithm for Forest Fire Prediction, 2023. 10.1007/978-981-99-4554-2_15.
16. Pande Chaitanya & Radwan Neyara & Heddham Salim & Othman Kaywan & Alshehri Fahad & Pal Subodh & Pramanik Malay. Forecasting of monthly air quality index and understanding the air pollution in the urban city, India based on machine learning models and cross-validation, *Journal of Atmospheric Chemistry*, 2024, 82, pp. 1-26. 10.1007/s10874-024-09466-x.
17. Salman Hasan & Kalakech Ali & Steiti Amani. Random Forest Algorithm Overview, *Babylonian Journal of Machine Learning*, 2024, pp. 69-79. 10.58496/BJML/2024/007.
18. Bijan Ahmed & Al-Rahim Ali. Random Forest and Decision Tree Facies Classification Models for Well Log Data of the Mishrif Formation from Basrah Oil Company, Southern Iraq, *Iraqi Geological Journal*, 2025, 57, pp. 14-32. 10.46717/igj.57.2E.2ms-2024-11-11.
19. Al-Janabia Samaher & AlShourbaji Ibrahim & Patel Ahmed. Applied Predicative Modeling to Improve Recommendation Systems for Forecasting of Fire Occurrences, 2015.
20. Ramadhan Rafiq & Ashari Wahid. Performance Comparison of Random Forest and Decision Tree Algorithms for Anomaly Detection in Networks, *Journal of Applied Informatics and Computing*, 2024, 8, pp. 367-375. 10.30871/jaic.v8i2.8492.
21. Li Liping. Comparative Research on Diabetes Influencing Factors Based on Random Forest and Decision Tree Models, *Highlights in Science, Engineering and Technology*, 2023, 72, pp. 231-242. 10.54097/7m4x7j04.
22. Hu Yaowen. Comparison and Analysis of the Effectiveness of Linear Regression, Decision Tree, and Random Forest Models for Health Insurance Premium Forecasting, *Advances in Economics, Management and Political Sciences*, 2024, 79, pp. 347-353. 10.54254/2754-1169/79/20241754.
23. Cansler C. & Wright Micah & Mantgem Phillip & Shearman Timothy & Varner J. & Hood Sharon. Drought before fire increases tree mortality after fire, *Ecosphere*, 2024, 15. 10.1002/ecs2.70083.

Сингх Санны – Южный федеральный университет, e-mail: singkh@sfedu.ru; г. Таганрог, Россия; тел.: +79885751350; кафедра систем автоматического управления; аспирант; ассистент.

Прибыльский Алексей Васильевич – Южный федеральный университет; e-mail: apribylsky@sfedu.ru; г. Таганрог, Россия; тел.: +79885619718; кафедра систем автоматического управления; к.т.н.; доцент.

Косенко Евгений Юрьевич – Южный федеральный университет; e-mail: ekosenko@sfedu.ru; г. Таганрог, Россия; тел.: +78634394474; кафедра систем автоматического управления; к.т.н.; доцент.

Singh Sanni – Southern Federal University; e-mail: singkh@sfedu.ru; Taganrog, Russia; phone: +79885751350; the Department of Automatic Control Systems; graduate student; assistant.

Pribylskiy Alexey Vasilievich – Southern Federal University; apribylsky@sfedu.ru; Taganrog, Russia; phone: +79885619718; the Department of Automatic Control Systems; cand. of eng. sc.; associate professor.

Kosenko Evgeny Yuryevich – Southern Federal University; ekosenko@sfedu.ru; Taganrog, Russia; phone: +78634394474; the Department of Automatic Control Systems; cand. of eng. sc.; associate professor.

УДК 519.863

DOI 10.18522/2311-3103-2025-1-81-91

Д.Ю. Зорькин, Л.В. Самофалова, Н.В. Асанова

МУРАВЬИНЫЙ АЛГОРИТМ НА PYTHON

Данное исследование посвящено анализу и оптимизации муравьиного алгоритма для решения задачи коммивояжера, являющейся классической NP-трудной проблемой комбинаторной оптимизации. Основная цель работы – экспериментальная оценка влияния параметров алгоритма на качество и эффективность поиска приближённых решений, а также разработка рекомендаций по их адаптивной настройке. В качестве тестового набора данных использован стандартный граф Berlin52 из библиотеки TSPLIB, содержащий координаты 52 городов с известным оптимальным маршрутом длиной