

ИСПОЛЬЗОВАНИЕ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ ДЛЯ ПРОГНОЗИРОВАНИЯ ПОЖАРОВ НА ОСНОВЕ МЕТЕОРОЛОГИЧЕСКИХ ДАННЫХ

В.Н. Палаткин¹, А.А. Андрияшин¹, Д.Д. Гетманцев¹

¹ФГБОУ ВО «Воронежский государственный лесотехнический университет
имени Г.Ф. Морозова»

Аннотация. Проблема лесных пожаров приобретает все большую актуальность в условиях изменения климата и увеличения частоты экстремальных погодных явлений. Прогнозирование пожаров на основе метеорологических данных, таких как температура, влажность, скорость ветра и осадки, играет важную роль в управлении рисками и минимизации ущерба. В данной статье рассматривается применение логистической регрессии (ЛР) для оценки вероятности возникновения лесных пожаров. ЛР является эффективным инструментом, обладающим преимуществами в плане интерпретируемости, скорости вычислений и гибкости в использовании различных типов данных. Несмотря на наличие более сложных моделей машинного обучения, таких как случайные леса, ЛР остается полезным методом для первичного анализа и моделирования. Результаты исследований показывают высокую точность предсказаний, что делает логистическую регрессию востребованным инструментом как в научных исследованиях, так и в практическом применении.

Ключевые слова: лесные пожары, прогнозирование пожаров, логистическая регрессия, метеорологические данные, машинное обучение, климатические изменения, управление рисками, температура, влажность, моделирование.

USING LOGISTIC REGRESSION FOR FIRE PREDICTION BASED ON METEOROLOGICAL DATA

V.N. Palatkin¹, A.A. Andryushin¹, D.D. Getmantsev¹

¹Voronezh State University of Forestry and Technologies named after G.F. Morozov

Abstract. The issue of forest fires is becoming increasingly relevant in the context of climate change and the growing frequency of extreme weather events. Fire prediction based on meteorological data, such as temperature, humidity, wind speed, and precipitation, plays a crucial role in risk

management and minimizing damage. This article discusses the application of logistic regression (LR) for assessing the probability of forest fires. LR is an effective tool with advantages in terms of interpretability, computational efficiency, and flexibility in handling various data types. Despite the existence of more complex machine learning models, such as random forests, LR remains a valuable method for initial data analysis and modeling. Research results demonstrate high prediction accuracy, making logistic regression a sought-after tool in both scientific research and practical applications.

Keywords: forest fires, fire prediction, logistic regression, meteorological data, machine learning, climate change, risk management, temperature, humidity, modeling.

Лесные пожары продолжают оставаться глобальной угрозой, особенно на фоне климатических изменений, усиливающих экстремальные погодные условия. Создание систем прогнозирования возгораний на основе метеорологических параметров — ключевой инструмент для предотвращения катастроф и снижения их последствий. Среди методов машинного обучения логистическая регрессия (ЛР) выделяется как эффективный подход для оценки вероятности возникновения пожаров через анализ взаимосвязей между факторами риска.

Логистическая регрессия (ЛР) является наиболее распространенным методом статистического анализа для моделирования бинарных исходов (0 или 1, да или нет). В отличие от линейной регрессии, суть которой заключается в предоставлении непрерывной зависимости, логистическая регрессия использует логистическую функцию для оценки вероятности наступления какого-либо события.

Логистическая регрессия имеет несколько преимуществ при прогнозировании лесных пожаров:

1. ЛР демонстрирует, как каждый фактор (например, температура или влажность) влияет на итоговый прогноз.
2. Алгоритм работает даже при ограниченном объеме данных и требует меньше вычислительных мощностей, чем нейросети или ансамблевые методы (например, Random Forest).[4]
3. Модель совместима с разными типами данных (категориальными, числовыми) и не требует их нормального распределения.
4. Простота архитектуры позволяет быстро настраивать и тестировать гипотезы.

Исследования показывают, что логистическая регрессия может достигнуть высокой точности в прогнозировании лесных пожаров. В одном из исследований было установлено, что модели ЛР обеспечивают точность предсказания до 81.6% для лесных пожаров и 70.4% для сельскохозяйственных[1][5].

Для построения моделей чаще всего используются:

1. Температура воздуха.
2. Относительная влажность.
3. Скорость и направление ветра.
4. Уровень осадков.
5. Индексы засухи (например, индекс Кейтча-Брайама).

Комбинация высокой температуры и низкой влажности, например, резко повышает риск возгорания [2]. Успешный пример применения ЛР — система прогнозирования в горных районах Мексики, где интеграция данных метеостанций и динамических индексов засухи значительно повысила точность предупреждений [2].

Другие исследования также показали эффективность ЛР в различных географических условиях. В качестве другого примера, в Китае была проведена работа по анализу уровня опасности лесных пожаров с использованием этой модели, где была достигнута высокая точность предсказаний[5].

Логистическая регрессия, при всей своей практичности, не является универсальным решением. Её главный недостаток — ограниченная способность выявлять сложные нелинейные взаимосвязи между переменными. Например, риск пожара может зависеть не только от линейной комбинации температуры и влажности, но и от их взаимодействия в определенных условиях (например, резкие перепады температур после засухи). В таких сценариях более эффективны следующие алгоритмы:

1. Случайный лес (Random Forest)

Этот метод строит ансамбль решающих деревьев, каждое из которых анализирует случайные подмножества данных и признаков. За счет этого он улавливает скрытые паттерны и нелинейные зависимости. В исследовании, посвященном прогнозированию пожаров в лесах Канады, Random Forest достиг точности 87.5%, что на 6% выше, чем у ЛР [1]. Однако за это приходится «платить»: модель требует значительных вычислительных ресурсов и времени для обучения, а её результаты сложнее интерпретировать из-за многослойной структуры.

2. Метод опорных векторов (SVM)

SVM ищет гиперплоскость, которая максимально разделяет классы в многомерном пространстве. Его сильная сторона — работа с высокоразмерными данными (например, спутниковые снимки в комбинации с метеопараметрами). В одном из экспериментов SVM продемонстрировал точность 84.3% в прогнозировании пожаров на территории Греции, но только после трудоемкой настройки

ядерных функций [3]. Для сравнения, ЛР не требует подобной оптимизации, что упрощает её внедрение.

3. Градиентный бустинг (Gradient Boosting)

Алгоритм последовательно улучшает предсказания за счет комбинации слабых моделей (обычно деревьев). Он особенно эффективен в задачах с несбалансированными данными (например, когда случаи пожаров редки). В исследовании для австралийских регионов градиентный бустинг превзошел ЛР на 12%, но при этом время обучения модели увеличилось в 5 раз [4].

Логистическая регрессия остается востребованным инструментом для оценки риска лесных пожаров. Её способность работать с разнородными данными, низкие требования к инфраструктуре и прозрачность результатов позволяют создавать базовые, но эффективные системы мониторинга. Хотя современные алгоритмы превосходят ЛР в сложных сценариях, она сохраняет статус «золотого стандарта» для стартовых исследований и оперативного реагирования.

Список литературы

1. Comparing machine learning algorithms to predict vegetation fire detections in Pakistan: сайт URL: <https://fireecology.springeropen.com/articles/10.1186/s42408-024-00289-5> (дата обращения 23.10.2024).
2. Daily prediction modeling of forest fire ignition using meteorological drought indices in the Mexican highlands: сайт URL: <https://iforest.sisef.org/contents/?id=ifor3623-014> (дата обращения 23.10.2024).
3. A Survey of the Machine Learning Models for Forest Fire Prediction and Detection: сайт URL: <https://www.scirp.org/journal/paperinformation?paperid=126584> (дата обращения 23.10.2024).
4. Logistic regression versus XGBoost for detecting burned areas using satellite images: сайт URL: <https://link.springer.com/article/10.1007/s10651-023-00590-7> (дата обращения 23.10.2024).
5. Применение модели логистической регрессии при принятии решений по определению количества привлекаемых сил на ликвидацию лесных пожаров: сайт URL: <https://www.fire-smi.ru/jour/article/view/1408> (дата обращения 23.10.2024).
6. Формализация верификации топологии и электрической схемы для систем автоматизированного проектирования / Т. В. Скворцова, К. В. Зольников, А. М. Плотников, И. В. Скоркин // Моделирование систем и процессов. – 2024. – Т. 17, № 3. – С. 61-70. – DOI 10.12737/2219-0767-2024-59-68. – EDN DUYQNJ.

7. Парамонов, А. А. Исследование структуры тропического циклона на основе анализа данных спутникового видеоряда с применением метода почти периодического анализа / А. А. Парамонов, А. В. Калач, А. С. Кравченко // Моделирование систем и процессов. – 2024. – Т. 17, № 4. – С. 67-76. – DOI 10.12737/2219-0767-2024-17-4-67-76. – EDN BOGSTY.

8. Асадов, Х. Г. Моделирование температурного поля при лесных пожарах / Х. Г. Асадов, Г. З. Байрамов // Моделирование систем и процессов. – 2024. – Т. 17, № 4. – С. 7-15. – DOI 10.12737/2219-0767-2024-17-4-7-15. – EDN ECSUWP.

9. Сумин, В. И. Особенности выбора членов экспертной группы для анализа функционирования сложной организационной системы силовых структур / В. И. Сумин, А. С. Дубровин, И. С. Куцева // Моделирование систем и процессов. – 2024. – Т. 17, № 4. – С. 77-83. – DOI 10.12737/2219-0767-2024-17-4-77-83. – EDN ISTAML.

10. Юрчишина, М. В. Алгоритмическая модель СППР «Оптимальный учебный план» / М. В. Юрчишина, К. И. Бушмелева // Моделирование систем и процессов. – 2024. – Т. 17, № 4. – С. 84-95. – DOI 10.12737/2219-0767-2024-17-4-84-95. – EDN OJFBGD.

11. Реализация оптимального построения комбинационного устройства и оценка надежности по выходному напряжению / Ф.В. Макаренко, А.С. Ягодкин, К.В. Зольников, О.А. Денисова // Моделирование систем и процессов. – 2021. – Т. 14, № 4. – С. 130-139. – DOI: 10.12737/2219-0767-2021-14-4-130-139.

References

1. Comparing machine learning algorithms to predict vegetation fire detections in Pakistan: URL: <https://fireecology.springeropen.com/articles/10.1186/s42408-024-00289-5> (accessed on 23.10.2024).

2. Daily prediction modeling of forest fire ignition using meteorological drought indices in the Mexican highlands: URL: <https://iforest.sisef.org/contents/?id=ifor3623-014> (accessed on 23.10.2024).

3. A Survey of the Machine Learning Models for Forest Fire Prediction and Detection: URL: <https://www.scirp.org/journal/paperinformation?paperid=126584> (accessed on 23.10.2024).

4. Logistic regression versus XGBoost for detecting burned areas using satellite images: URL: <https://link.springer.com/article/10.1007/s10651-023-00590-7> (accessed on 23.10.2024).

5. Application of the logistic regression model for decision-making in determining the number of forces to combat forest fires: URL: <https://www.fire-smi.ru/jour/article/view/1408> (accessed on 23.10.2024).
6. Formalization of verification of topology and electrical circuit for computer-aided design systems / T. V. Skvortsova, K. V. Zolnikov, A. M. Plotnikov, I. V. Skorkin // Modeling of systems and processes. – 2024. – Vol. 17, No. 3. – Pp. 61-70. – DOI 10.12737/2219-0767-2024-59-68. – EDN DUYQHJ.
7. Paramonov, A. A. Study of the structure of a tropical cyclone based on the analysis of satellite video data using the method of almost periodic analysis / A. A. Paramonov, A. V. Kalach, A. S. Kravchenko // Modeling of systems and processes. – 2024. – Vol. 17, No. 4. – Pp. 67-76. – DOI 10.12737/2219-0767-2024-17-4-67-76. – EDN BOGSTY.
8. Asadov, H. G. Modeling of the temperature field during forest fires / H. G. Asadov, G. Z. Bayramov // Modeling of systems and processes. – 2024. – Vol. 17, No. 4. – P. 7-15. – DOI 10.12737/2219-0767-2024-17-4-7-15. – EDN ECSUWP.
9. Sumin, V. I. Features of the selection of members of the expert group for the analysis of the functioning of a complex organizational system of law enforcement agencies / V. I. Sumin, A. S. Dubrovin, I. S. Kushcheva // Modeling of systems and processes. – 2024. – Vol. 17, No. 4. – pp. 77-83. – DOI 10.12737/2219-0767-2024-17-4-77-83. – EDN ISTAML.
10. Yurchishina, M. V. Algorithmic model of the DSS "Optimal curriculum" / M. V. Yurchishina, K. I. Bushmeleva // Modeling of systems and processes. – 2024. – Vol. 17, No. 4. – pp. 84-95. – DOI 10.12737/2219-0767-2024-17-4-84-95. – EDN OJFBGD.
11. Realization of the optimal construction of a combination device and evaluation of reliability by output voltage / Ph.V. Makarenko, A.S. Yagodkin, K.V. Zolnikov, O.A. Denisova // Modeling of systems and processes. – 2021. – Vol. 14, No. 4. – pp. 130-139. – DOI: 10.12737/2219-0767-2021-14-4-130-139.