

СПИСОК ЛИТЕРАТУРЫ

1. Kuznetsov G.V., Kravchenko E.V. Influence of the air gap between two cells of the storage battery on the thermal conditions of its operation: Numerical analysis and reliability assessment // Journal of Electroanalytical Chemistry. – 2023. – Vol. 945. – P. 117688. DOI 10.1016/j.jelechem.2023.117688. – EDN WKKPZA.
2. Kuznetsov G.V., Kravchenko E.V. Estimation of the representative temperatures of a typical electrochemical battery cell when cooling under natural convection conditions // Journal of Energy Storage. – 2023. – Vol. 73. – P. 109168. DOI 10.1016/j.est.2023.109168. – EDN NGYIGZ.

ПРЕДСКАЗАНИЕ ЛЕСНЫХ ПОЖАРОВ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Е.В. Гузеев

Томский политехнический университет, evg39@tpu.ru

Научный руководитель: Н.В. Барановский, к.ф.-м.н., доцент, ТПУ

Лесные пожары – одна из наиболее разрушительных и труднопрогнозируемых катастроф, оказывающая существенное воздействие на экосистему, экономику и безопасность человека. Проблема прогнозирования лесных пожаров является чрезвычайно актуальной, поскольку своевременное обнаружение и эффективное реагирование могут значительно снизить масштабы и последствия пожаров. Традиционные методы прогнозирования, основанные на статистических моделях и эмпирических данных, часто оказываются недостаточно точными и оперативными для принятия своевременных мер.

В последние годы методы машинного обучения продемонстрировали высокую эффективность в решении задач анализа данных, классификации и прогнозирования. Применение машинного обучения в области прогнозирования лесных пожаров открывает новые возможности для повышения точности и оперативности прогнозов, что, в свою очередь, позволит оптимизировать деятельность пожарных служб и снизить риски для населения и окружающей среды.

В данной статье рассматривается актуальность проблемы прогнозирования лесных пожаров и предлагается использование методов машинного обучения в качестве эффективного инструмента для решения данной задачи. В работе анализируются существующие подходы к прогнозированию лесных пожаров, рассматриваются преимущества и недостатки различных методов МО, а также обсуждаются возможные направления дальнейших исследований в данной области.

Целью исследования является сравнение моделей для прогнозирования пожаров для заиграевского района республики Бурятия.

Обзор

Во Gao и др. [1] проводили исследование по сравнению моделей прогнозирования лесных пожаров в провинции Цзилинь, Китай: логистической регрессии (LR), географически взвешенной логистической регрессии (GWLRL), лассо-регрессии (Lasso), случайного леса (RF) и метод опорных векторов (SVM). Случайный лес показал наивысшую точность (88.7 %), за ним следует SVM (87.7 %). GWLRL и Lasso достигли точности 86.0 и 85.0 % соответственно, в то время как LR показала самую низкую точность (84.6 %). Основными факторами, влияющими на вероятность возникновения пожаров, являются климатические переменные (среднее атмосферное давление, температура, давление водяного пара, минимальная относитель-

ная влажность), а также топографические факторы (высота над уровнем моря). Социально-экономические факторы (плотность населения, ВРП) влияют в меньшей степени. Пространственный анализ показал, что наибольшая вероятность пожаров наблюдается в юго-восточной части провинции, где преобладают леса и горы, в то время как в равнинных районах запада вероятность ниже. Исследование подтверждает эффективность алгоритмов машинного обучения, таких как RF и SVM, для прогнозирования лесных пожаров, однако для практического применения могут быть более подходящими модели, сочетающие высокую точность с относительной простотой реализации, такие как GWLR.

Для создания эффективных моделей прогнозирования лесных пожаров, как показано в исследованиях [1] и [2], крайне важно получение качественных и репрезентативных данных. В частности, данные о точках отсутствия пожара играют ключевую роль в обучении моделей, поскольку они позволяют модели учиться различать области, где пожары маловероятны, от тех, где они могут возникнуть.

В статье [2] был предложен метод сбора данных о точках отсутствия пожара на основе географического сходства. Этот метод позволяет создавать более репрезентативные наборы данных, что, в свою очередь, повышает точность моделей машинного обучения. Исследование показало, что модели, обученные на данных, собранных с использованием предложенного метода, демонстрируют значительно более высокую точность моделирования и прогнозирования по сравнению с традиционными методами.

В последние годы значительное внимание уделяется применению глубокого обучения для прогнозирования лесных пожаров. Например, Фан и др. [3] разработали легковесную модель обнаружения лесных пожаров на основе глубокого обучения. Сатхишкumar и др. [4] использовали глубокое обучение для обнаружения дыма от пожаров. Канг и др. [5] применили глубокое обучение с использованием геостационарных спутниковых данных для обнаружения пожаров с сокращенным временем задержки.

Однако, несмотря на успехи в области машинного и глубокого обучения, существует ряд проблем, связанных с прогнозированием лесных пожаров. В частности, сложность заключается в учете пространственно-временных закономерностей и периодичности факторов, влияющих на возникновение пожаров. В статье [6] предлагается использовать модель LSTNet для прогнозирования лесных пожаров, которая способна эффективно учитывать как краткосрочные, так и долгосрочные зависимости между факторами, влияющими на пожары.

Результаты исследования

Предсказание пожаров проводилось для взятых тестовых данных, которые состояли из пожаров и смещенных точек не пожаров.

Данные:

Были собраны данные о погодных условиях, рельефа, осадках, населении и исторических данных о пожарах с 2003.

Данные были очищены от пропусков и аномалий, а также преобразованы в подходящий для моделирования формат.

Разбиение данных:

Для обучения моделей данные были разбиты на обучающую и валидационную выборку (30 % валидационной выборки, 70 % обучающей).

Применялась стратифицированная выборка для сохранения баланса классов (наличие/отсутствие пожара).

Используемые алгоритмы:

Для предсказания использовались алгоритмы градиентного бустинга: CatBoost, XGBoost, LightGBM.

Также были протестированы алгоритмы случайного леса (RandomForest), метода опорных векторов (SVM) и логистической регрессии (Logistic Regression).

Ключевая метрика:

В качестве ключевой метрики использовалась метрика accuracy, так как данные сбалансированы.

Результаты

Результаты сравнения качества работы алгоритмов приведены в табл. 1.

Таблица 1. Результаты сравнения качества работы алгоритмов

Алгоритм	Метрика <i>accuracy</i>
CatBoost	0,97239
LightGBM	0,98303
XGBoost	0,945
RandomForest	0,91796
SVM	0,793
Логистическая регрессия	0,669

Заключение

В результате проведенного исследования было установлено, что наиболее успешно с задачей предсказания пожаров справились алгоритмы градиентного бустинга, а именно LightGBM и CatBoost, которые продемонстрировали высокую точность (accuracy) – 0,98303 и 0,97239 соответственно. Эти результаты значительно превосходят показатели других протестированных алгоритмов, таких как RandomForest, SVM и логистическая регрессия. Рекомендуется использовать модель LightGBM для практического применения в системе раннего предупреждения о возможности возникновения пожаров. Для дальнейшего повышения точности предсказаний можно рассмотреть возможность построения ансамбля моделей, увеличения объема данных и включения дополнительных признаков, таких как данные с метеорологических спутников и информация о деятельности человека в лесу. Результаты исследования демонстрируют высокую эффективность использования алгоритмов градиентного бустинга для решения задачи предсказания пожаров, что открывает перспективы для создания системы раннего предупреждения и снижения рисков, связанных с лесными пожарами.

СПИСОК ЛИТЕРАТУРЫ

1. Bo Gao, Yanlong Shan, Xiangyu Liu, Sainan Yin, Bo Yu, Chenxi Cui & Lili Cao. Prediction and driving factors of forest fire occurrence in Jilin Province. – China, 2023.
2. Xu Q., Li W., Liu J., Wang X. A geographical similarity-based sampling method of non-fire point data for spatial prediction of forest fires // Forest Ecosystems. – 2023. – V. 10(1). – P. 100104.
3. Fan R., Pei M. Lightweight Forest Fire Detection Based on Deep Learning // In Proceedings of the 2021 IEEE 31st International/ Workshop on Machine Learning for Signal Processing (MLSP), Gold Coast, QLD, Australia, 25–28 October 2021; pp. 1–6.
4. Sathishkumar V.E., Cho J., Subramanian M., Naren O.S. Forest fire and smoke detection using deep learning-based learning without forgetting // Fire Ecol. – 2023. – V. 19. – P. 9.
5. Kang Y., Jang E., Im, J., Kwon C.G. A deep learning model using geostationary satellite data for forest fire detection with reduced detection latency // GISci. Remote Sens. – 2022. – V. 59. – P. 2019–2035.
6. Lin X., Li Z., Chen W., Sun X., Gao D. Forest Fire Prediction Based on Long- and Short-Term Time-Series Network. Forests. – 2023. – 14(4). – P. 778. <https://doi.org/10.3390/f14040778>