

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В СФЕРЕ ПОЖАРНОЙ БЕЗОПАСНОСТИ, МАШИННОЕ ОБУЧЕНИЕ КЛАССИФИКАЦИИ РИСКА ВОЗНИКНОВЕНИЯ ПОЖАРА В ЗАДАННОЙ ОБЛАСТИ

Сергеева А. О., студент,

Научный руководитель: Портнов К. В., к. т. н., доцент,

Самарский государственный технический университет, г. Самара, Россия.

Аннотация: Цель исследования – разработать модель классификации риска возникновения пожара в заданной области. Статья обобщает практический опыт обучения модели на основе анализа данных о регионе, солнечной радиации, количестве осадков, температуре воздуха, относительной влажности. В результате исследования использовался датасет «California Environmental Conditions Dataset», который содержит данные наблюдений. При разработке происходила оценка качества модели, что показало возможность её применения в сфере безопасности.

Ключевые слова: машинное обучение, датасет, модель, классификация, разработка.

Природные явления, а именно пожары, одни из наиболее серьёзных угроз для безопасности людей и окружающей среды, особенно в районах засухи и высокой температуры воздуха. Прогнозирование и классификация риска возникновения пожара в заданной области играют важнейшую роль в предотвращении катастроф, минимизации ущерба и своевременном принятии необходимых мер.

Современные технологии меняют привычный уклад повседневной жизни, сфера пожарной безопасности также претерпевает изменения. Искусственный интеллект (ИИ) сегодня занимает центральное место в разработке новых решений, направленных на повышение уровня защиты от пожаров [5]. Эти

технологии позволяют не только оперативно реагировать на чрезвычайные ситуации, но и предотвращать их благодаря анализу данных и прогнозированию рисков [2]. Искусственный интеллект применяют в камерах слежения, за счет чего уменьшается время обнаружения пожара, в многофакторном анализе, где происходит анализ данных с датчиков. Интеграция искусственного интеллекта минимизирует ущерб и позволяет обнаружить возгорание на ранних стадиях.

Внедрение искусственного интеллекта в сферу пожарной безопасности основано на способности обрабатывать большие объемы данных, выявлять закономерности и предсказывать риски [1, 8].

Для машинного обучения классификации риска возникновения пожара модели был выбран открытый датасет «California Environmental Conditions Dataset» с платформы kaggle. Датасет включает в себя информацию о климатических условиях (15 признаков), где target – целевая бинарная переменная, показывающая риск возникновения пожара (0 – низкая вероятность возникновения пожара, 1 – высокая вероятность) [3].

В исследовании используются модели классификаторов, такие как методы опорных векторов с линейным ядром (Linear SVM), с полиномиальным ядром (Polynomial Kernel SVM) , с радиально-базисным ядром (RBF SVM), метод k–ближайших соседей (kNN), дерево решений (Decision tree), случайный лес (Random forest), градиентный бустинг (Gradient boosting), а также гауссовский наивный байесовский классификатор (GaussianNB) [9].

Для того, чтобы подготовить модель к обучению следует совершить предобработку данных в несколько этапов:

- Удаление лишней информации из датасета;
- Разделение датасета на входные и выходные данные;
- Разделение датасета на обучающую и тестовую выборку [4].

После выполнения всех вышеперечисленных этапов, которые подготавливают модель для машинного обучения, происходит оценка качества моделей в Таблице 1, с использованием метрики точности.

Таблица 1 – Точность моделей

Модель	Точность (%)
Linear SVM	95,3
Polynomial Kernel SVM	94,5
RBF SVM	95,1
kNN	93,7
Decision tree	94,3
Random forest	96,2
Gradient boosting	95,8
GaussianNB	92,2

По результатам сравнения наибольшую точность показала модель случайный лес (Random forest), её точность – 96,2%. После применения нормализации точность модели составила 96,1%, метод главных компонент с нормализацией – 96,1%, метод главных компонент без нормализации – 96,0%. Полученные данные показывают, что после применения методов значительно уменьшается количество признаков для обучения, увеличивается скорость обработки, но при этом снижается метрика точности. Оптимальным количеством признаков для дисперсии МГК с нормализацией является 10, для МГК без нормализации – 4 [6, 7].

На Рисунке 1 и 2 показаны дисперсии при изменении количества признаков для МГК с нормализацией и МГК без нормализации.

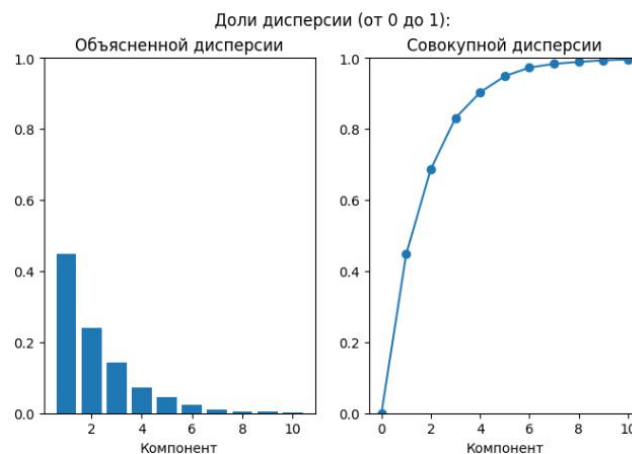


Рисунок 1 – Результаты дисперсий для МГК с нормализацией

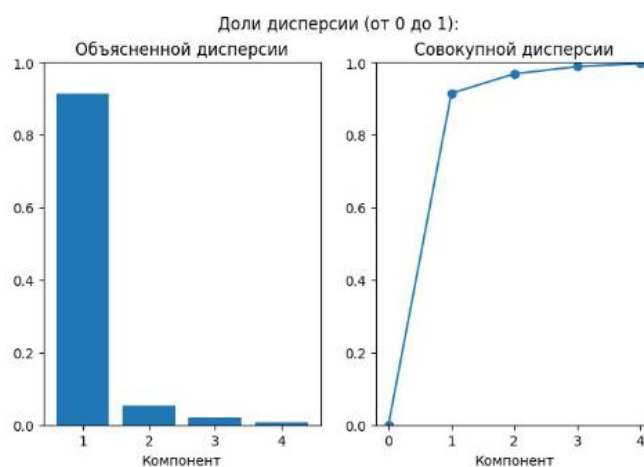


Рисунок 2 – Результаты дисперсий для МГК без нормализации

В данной статье была приведена разработка модели классификации риска возникновения пожара в заданной области, которая показала, что искусственный интеллект достаточно точно работает на обучающих данных. По этой причине его использование в реальной системе раннего предупреждения о пожарах в сфере пожарной безопасности поможет предотвратить возгорания на начальных этапах, а также способен обеспечить автоматизацию обнаружения.

Литература

1. Забержинский, Б. Э. Разработка алгоритма модели машинного обучения для детектирования пневмоний на рентгенограммах // Экономика и качество систем связи. – 2024. – № 4(34). – С. 144-152. – EDN HHRQXG.
2. Золин, А.Г. Модель прогнозирования паттернов поведения индивида при обработке видеоизображений // Прикаспийский журнал: управление и высокие технологии. – 2024. – № 4(68). – С. 68-74.
3. Каримов, Б. Ф. Проблемы адаптации генетических алгоритмов к решению задач структурно-параметрической оптимизации // Современные исследования: теория, практика, результаты : Сборник материалов Международной научно-практической конференции, Москва, 29 декабря 2023 года. – Москва: Центр развития образования и науки, ООО "Издательство АЛЕФ", 2023. – С. 443-448. – DOI 10.26118/1590.2023.63.10.010. – EDN PUGPKI.
4. Ларкина, А. А. Алгоритм формирования обучающей выборки на основе метода кластеризации // Журнал монетарной экономики и менеджмента. – 2024. – № 6. – С. 38-42. – DOI 10.26118/2782-4586.2024.87.62.004. – EDN BHZKRR.
5. Олин, Р. А. Инновационный менеджмент предприятия в условиях взаимодействия с машинными клиентами и автономными агентами на основе искусственного интеллекта // Журнал монетарной экономики и менеджмента. – 2024. – № 3. – С. 218-224. – DOI 10.26118/2782-4586.2024.24.43.033. – EDN OJZVMH.
6. Портнов, К. В. Генетические алгоритмы и поиск эффективных порядков индикаторов в биржевой торговой стратегии на основе пересечения трех скользящих средних // Вестник Самарского государственного технического университета. Серия: Технические науки. – 2005. – № 32. – С. 72-76. – EDN JWUXKZ.
7. Портнов, К. В. Модификация модели генетических алгоритмов с применением нечеткого оператора выбора точки Кроссовера / К. В. Портнов //

Математическое моделирование и краевые задачи : Труды Второй Всероссийской научной конференции, Самара, 01–03 июня 2005 года / Редколлегия: В. П. Радченко (отв. редактор), Э. Я. Рапопорт, Е. Н. Огородников, М. Н. Саушкин (отв. секретарь). Том Часть 2. – Самара: Самарский государственный технический университет, 2005. – С. 203-205. – EDN TGCFVB.

8. Фошин, М.А. Синтез программного алгоритма и архитектуры приложения по автоматизированному сбору информации в сетевых агрегаторах // Экономика и качество систем связи. – 2024. – № 3(33). – С. 97-108. – EDN LKXEXG.

9. Харитонов, Е.А. Разработка автоматизированной информационной системы скоринга на основе многофакторной логистической регрессии // Журнал монетарной экономики и менеджмента. – 2024. – № 3. – С. 157-167. – DOI 10.26118/2782-4586.2024.23.62.024. – EDN SEEWTV.