



АКАДЕМИЯ
ГРАЖДАНСКОЙ ЗАЩИТЫ
МЧС РОССИИ

«Применение технологий обработки больших данных для построения моделей прогнозирования лесных (ландшафтных) пожаров»

Докладчик: Иванов Евгений Вячеславович, кандидат технических наук, начальник научно-исследовательского отдела (по проблемам ГО и ЧС)

Содокладчик: Дмитриев Антон Викторович, научный сотрудник научно-исследовательского отдела (по проблемам ГО и ЧС)

Химки 2022

Цель и задачи работы

Цель:

Повышение эффективности прогнозирования последствий природных пожаров на основе данных дистанционного зондирования Земли за счёт моделирования и алгоритмизации процессов прогнозирования распространения пожаров.

Задачи:

разработать алгоритм пространственной группировки природных пожаров с помощью методов кластеризации;

сформировать модель оценки вероятности возникновения природных пожаров в заданной географической области на основе обработанных данных дистанционного зондирования Земли и метеорологических условий;

разработать алгоритм для программной реализации в информационной системе космического мониторинга чрезвычайных ситуаций МЧС России.

Система исходных данных

Количество осадков на рассматриваемом участке местности

- По данным сети станций метеорологического наблюдения

Количество термоточек (событий) на рассматриваемом участке местности

- По данным спутниковой группировки

Индекс вегетационной активности

- По данным спутниковой группировки

Плотность дорожной сети, число населенных пунктов

- Статистические справочники

Сведения по температуре воздуха, солнечной активности

- По данным сети станций метеорологического наблюдения

Гипотезы

Гипотеза 1.

Для определенной территории можно выделить такую совокупность признаков пожарной опасности (различной природы) и их значений, которые будут соответствовать условиям возникновения пожаров.

Гипотеза 2.

Для определенной территории условия возникновения пожаров не изменяются в каждый пожароопасный сезон.

Гипотеза 3.

Условия возникновения пожаров позволяют оценить уровень природной пожарной опасности.

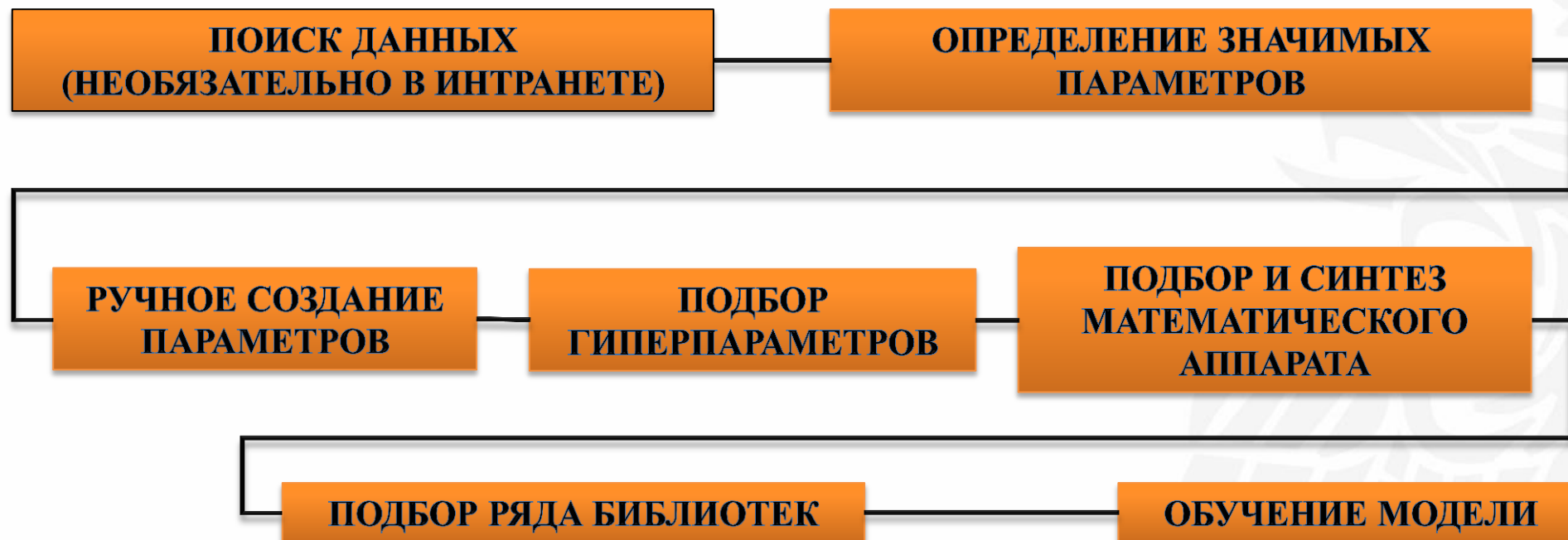
Гипотеза 4.

Для территорий с более высоким уровнем природной пожарной опасности пожары наступают либо раньше, либо чаще.

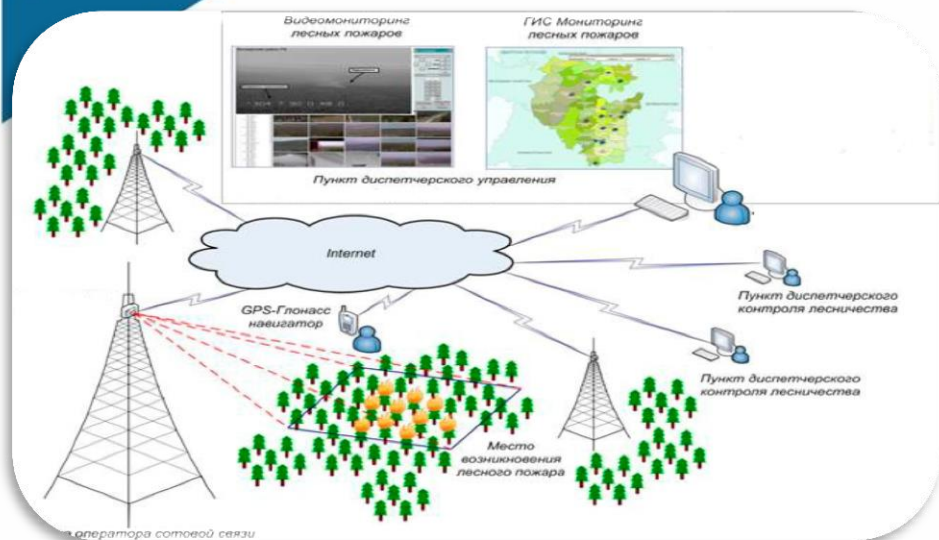
Гипотеза 5.

Существуют предпосылки (предвестники) чрезвычайных лесопожарных ситуаций.

Разработка моделей



Система наблюдений и мониторинга



- Температура (airc)
- Влажность (rhum)
- Скорость ветра верт. (wwnd)
- Скорость ветра гор. (vwnd)
- Давление (hgt)

* Параметры выбраны на основании существующих прогнозируемых



Модели обработки данных

МОДЕЛЬ FEDOT

Входные
данные

XGBReg

XGBReg
предсказание

Scaling

PCA

PCA
предсказание

Вывод
окончательного
прогноза

МОДЕЛЬ LAMA

Входные
данные

LightGBM
MOD 0

LightGBM
предсказание

CatBoost
MOD 2

CatBoost
предсказание

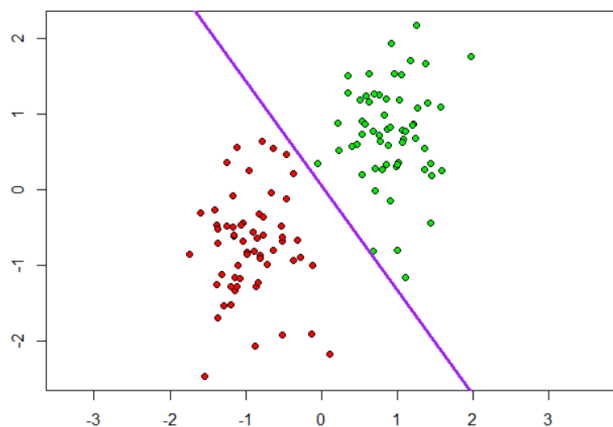
CatBoost
MOD 3

CatBoost
предсказание

Средневзвешенное
значение

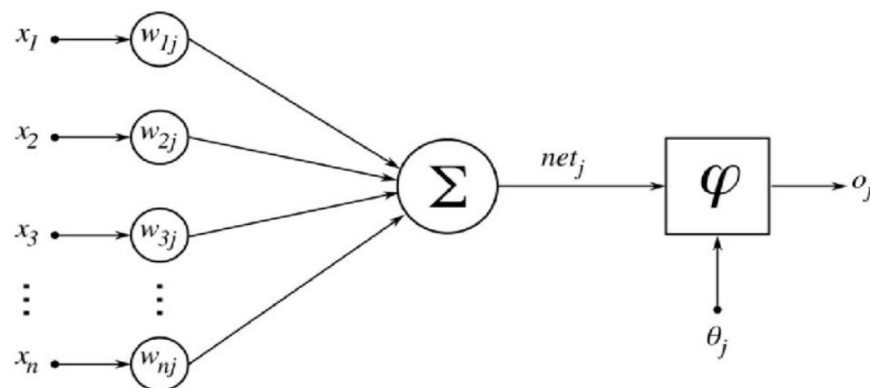
Вывод
окончательного
прогноза

Сравнительные результаты моделей



FEDOT

- Точность на обработанных данных (98%)
- Точность предсказания на стандартной выборке (99%)
- Точность на выборке за месяц (36%)
- Масштаб (2.5° на 2.5° широты и долготы соответственно)



LAMA

- Точность предсказания на день (86%)
- Точность предсказания на стандартной выборке (93%)
- Точность на выборке за месяц (85%)
- Радиус предсказания (0.625° на 0.625° широты и долготы соответственно)

Озеро данных МЧС России

Amudsen

Superset

Jupyter

На этой платформе мы получаем доступ к физической модели баз данных

Современная платформа, выполняющая анализ и визуализацию данных

Выполняется анализ, обработка данных, построение моделей, обучение ИИ

Исторические данные

Текущие данные

Прогнозные данные

Основные результаты

Как предсказываются данные на 2021 год

Правильно предсказываются **не пожары** в 53845 из 61554 случаев (87.5%)

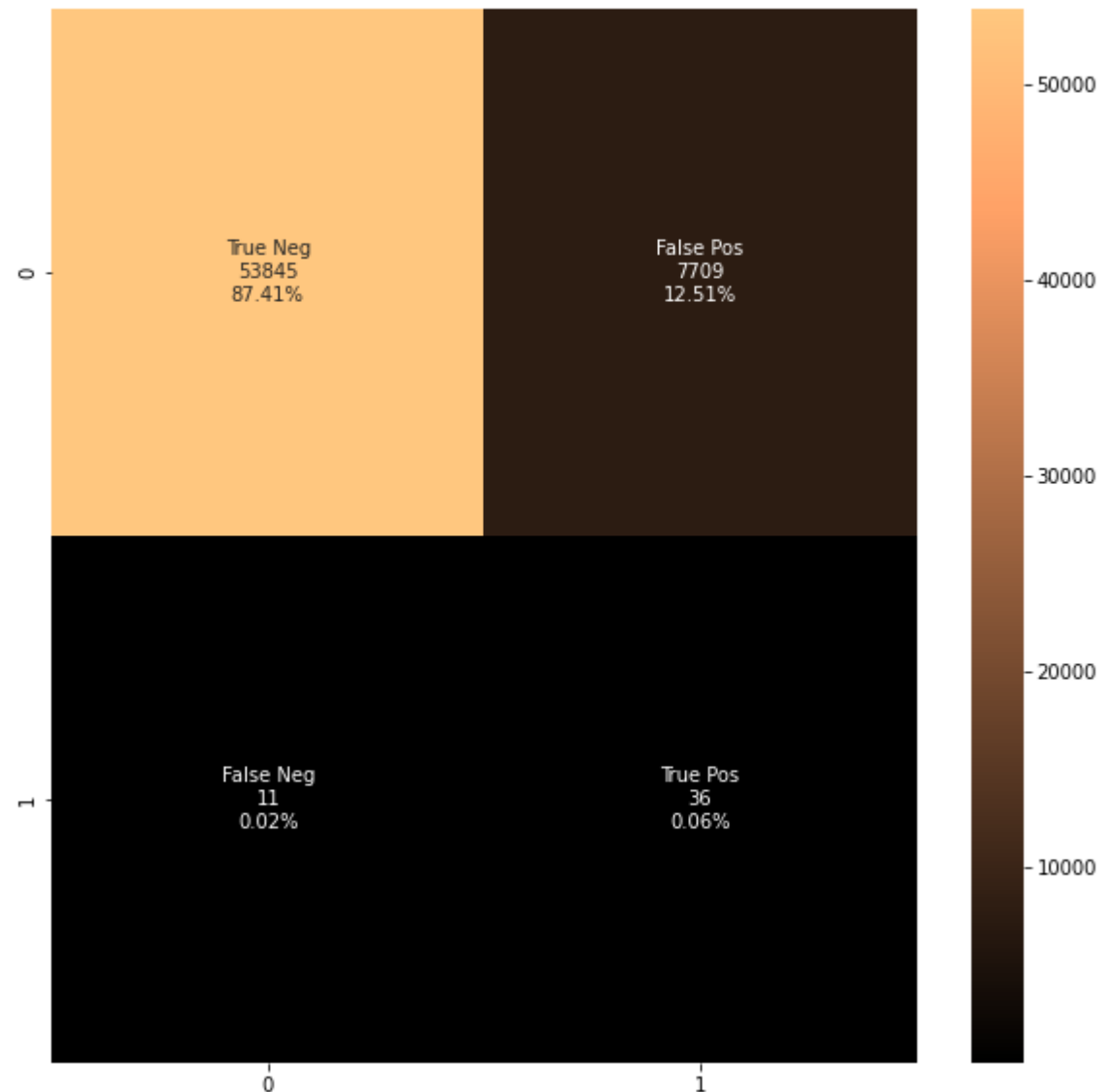
Правильно предсказываются **пожары** в 36 из 47 случаев (76.6 %)

Ошибочно предсказываются **пожары** в 11 из 36 случаев (23.4 %)

Ошибочно предсказываются **не пожары** в 7709 из 61554 случаев (12.5%)

Test 2021

<matplotlib.axes._subplots.AxesSubplot at 0x7f36341a7490>



Основные результаты

Как предсказываются данные на 2021 год с 5 по 9 месяц

Правильно предсказываются **не пожары** в 26365 из 34074 случаев (77.4 %)

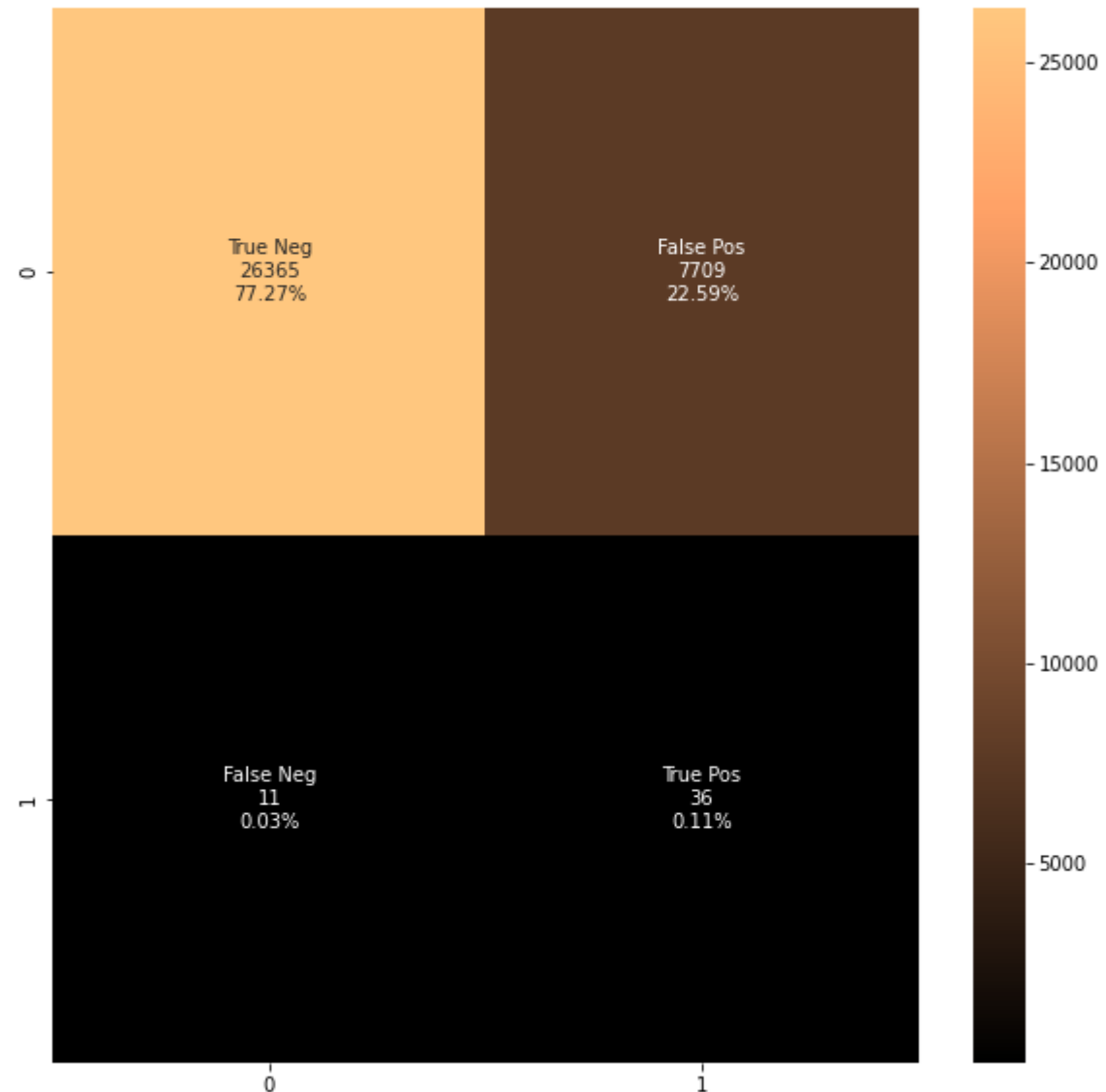
Правильно предсказываются **пожары** в 36 из 47 случаев (76.6 %)

Ошибочно предсказываются **пожары** в 11 из 36 случаев (23.4 %)

Ошибочно предсказываются **не пожары** в 7709 из 34074 случаев (22.6 %)

Test 2021 5-9

<matplotlib.axes._subplots.AxesSubplot at 0x7f36340b16d0>



Выводы по работе

На имеющемся наборе данных получение точного предсказания пожаров проблематично, что видно по корреляции значений параметров с метками пожаров;

Модель может иметь практическую значимость, так как она отсеивает большинство случаев, не представляющих угрозы. Модель не дает точного предсказания пожара (так как по видимому процесс их возникновения не детерминирован исследуемыми факторами), позволяет сосредоточить внимание на более пожароопасных зонах.



Благодарю за внимание!

Иванов Евгений Вячеславович