

ОНЛАЙН-ХАКАТОН

EMERGENCY DATAHACK

ТРЕК 3



МЧС
РОССИИ



ИНИД



ТРЕК 3

Предсказание вероятности возникновения ДТП

ОБЩЕЕ ОПИСАНИЕ ЗАДАЧИ

Построить предиктивную модель, которая позволит на основании данных о погодных условиях и интенсивности движения спрогнозировать возникновение аварии и ее тяжесть на определенном участке федеральной автомобильной дороги на конкретную дату и время.

ЗАЧЕМ МЫ ЭТО ДЕЛАЕМ?

Национальный проект «Безопасные и качественные автомобильные дороги» к 2024 году предполагает постепенное снижение количества погибших в ДТП и мест концентрации ДТП.

Федеральные автомобильные дороги являются важнейшими с точки зрения экономики и национальной безопасности транспортными узлами Российской Федерации, но в то же время возрастают и риски возникновения ДТП на всей протяженности трасс с такого рода высокой эксплуатационной нагрузкой и интенсивностью дорожного трафика.

Способ реагирования МЧС России на ДТП в том числе зависит от того является ли ДТП чрезвычайной ситуацией. Режим чрезвычайной ситуации объявляется от 5 погибших и/или 10 пострадавших, поэтому важно предсказывать не только факт возникновения ДТП, но и его соответствующую категорию.

Решение такого рода задачи позволит реализовать МЧС России следующие мероприятия:

расчет оптимально необходимого количества сил и средств пожарно-спасательных подразделений в определенные временные отрезки мониторинга за ситуацией с безопасностью на дороге;
оценка необходимого наращивания сил и средств в зависимости от потенциального уровня тяжести ДТП.

ФОРМУЛИРОВКА ЗАДАЧИ

Участникам предстоит разработать алгоритм прогнозирующий наиболее вероятную категорию события дорожно-транспортного происшествия (далее по тексту «ДТП»):

ДТП не произойдет;

ДТП произойдет без пострадавших;

ДТП произойдет с пострадавшими;

Режим ЧС (объявляется от 5 погибших и/или 10 пострадавших).

МЕТРИКА

У предиктивной модели предполагается 1 таргет-переменная: у нее может быть значение 0,1,2,3 (0 — нет аварии / 1 — есть авария, без пострадавших / 2 — есть авария, с пострадавшими / 3 — ЧС). Для оценки алгоритма будет использоваться macro-F1 score.

ОЖИДАЕМЫЕ РЕЗУЛЬТАТЫ

Краткая аннотация проекта

описание проекта в 2-4 предложениях;

ML-модель, обученная на тренировочной выборке

Ссылка на я.диск/dropbox с ML-моделью, обученной на тренировочной выборке;

Код, использованный при подготовке модели

Ссылка на код, использованный при подготовке модели (ссылка на любой открытый репозиторий с кодом, с помощью которого строилась модель, а также кодом, запускающим модель).

Презентация результатов решения

Ссылка на презентацию для очной защиты

КРИТЕРИИ ОЦЕНКИ

1. **60 баллов** распределяется в соответствии с **результатом, который ML-модель показала на соревновании**. Будет учитываться не место в турнирной таблице напрямую, а логарифмированное значение ошибки на тестовой выборке относительно величины ошибок других участников. По результатам соревнования на очную защиту перед жюри допускается не более 10 команд.

40 баллов распределяется по результатам оценки кода, выложенного в открытый репозиторий,
а также выступления на очной защите в соответствии со следующими критериями:

2. обоснованность выбранных методов для построения модели (**20 баллов**);
3. оригинальность подхода, в частности, использование дополнительных данных из открытых источников для повышения качества модели (**10 баллов**);
4. качество кода: его чистота и понятность (**5 баллов**);
5. качество очной защиты: последовательность и понятность изложения (**5 баллов**).

ИНСТРУКЦИЯ ПО ПЛАТФОРМЕ

- Перейти по ссылке <https://apply.emergencydatahack.ru/competition> или выбрать раздел “DS Чемпионат”
- Выбрать трек, в котором вы участвуете
- Нажать на кнопку “Скачать датасет”. Пароль для доступа к данным: **gang7leN**

The screenshot displays the 'DS ЧЕМПИОНАТ' (DS Championship) page on the Emergency Data Hack platform. On the left is a purple sidebar with navigation links: 'Панель управления', 'Новости', 'Анкета', 'Команда', 'Решение', 'DS Чемпионат' (highlighted with a red box and an arrow), 'Подтверждение', and 'Выйти'. The main content area has a header 'DS ЧЕМПИОНАТ' with a red arrow pointing to a track selection bar containing 'Трек 1', 'Трек 2', and 'Трек 3'. Below this, the selected track is 'ТРЕК #1 ПРЕДСКАЗАНИЕ ЗАТОРОВ НА РЕКЕ ЛЕНА В ПАВОДКООПАСНЫЙ ПЕРИОД'. The 'ОПИСАНИЕ ЗАДАЧИ' (Task Description) section explains the goal: to build a predictive model for ice jams on the Lena river. A purple button 'Скачать датасет' (Download dataset) is present. The 'РЕШЕНИЯ ДЛЯ ТЕСТОВОГО ДАТАСЕТА' (Solutions for test dataset) section includes instructions on file format and submission limits. At the bottom, a table shows submission history.

Код	Дата	Скор	*.CSV	*.ZIP
-2	23.05.21 23:55:11	0	+	+

Codebook доступен
по [ссылке](#).

ИНСТРУКЦИЯ ПО ПЛАТФОРМЕ

- Прочитать все, что написано на странице
- Для отправки решения выбрать CSV файл с ответами и архив с кодом, нажать кнопку “отправить” - необходимо отправлять архив и CSV каждый раз, чтобы жюри смогли провалидировать способ, которым был получен скор
- Для просмотра текущего лидерборда - нажмите кнопку “открыть LeaderBoard”

The screenshot shows the EMERGENCY DATANACK platform interface. On the left is a purple sidebar with navigation links: RU, EN, EMERGENCY DATANACK, ОНЛАЙН-КАКАТОН, Панель управления, Новости, Анкета, Команда, Решение, DS Чемпионат, Подтверждение, and Выйти. The main content area has a header "ОПИСАНИЕ ЗАДАЧИ" and a description of the task: "Построить предиктивную модель, которая позволит за 60 дней рассчитать вероятность возникновения заторов льда на каждые сутки в период ледохода на нескольких участках реки Лена." Below this is a purple button "Скачать датасет".

The next section is "РЕШЕНИЯ ДЛЯ ТЕСТОВОГО ДАТАСЕТА". It contains text about uploading solutions and a table of attempts. Below the table are two input fields for uploading files, each with a red box and an arrow pointing to it. At the bottom, there is a purple button "Отправить" and a black button "Открыть LeaderBoard" with a red box around it.

ОПИСАНИЕ ЗАДАЧИ

Построить предиктивную модель, которая позволит за 60 дней рассчитать вероятность возникновения заторов льда на каждые сутки в период ледохода на нескольких участках реки Лена.
[Подробнее описание.](#)

Скачать датасет

РЕШЕНИЯ ДЛЯ ТЕСТОВОГО ДАТАСЕТА

Загрузи решения для тестового датасета. Пример, как должен выглядеть файл для загрузки, коды ошибок и метрика: [подробнее](#). У вас есть 35 попыток на время всего соревнования, в зачет идет лучшая. Если возвращается ошибка - попытка не сгорает.

Попыток осталось: 34/35

История попыток:

Код	Дата	Скор	*.CSV	*.ZIP
-2	23.05.21 23:55:11	0		
0	23.05.21 23:55:27	0.006783842121492523		

Загрузить (*.CSV, лимит 3mb)

Загрузить (*.ZIP, лимит 3mb)

Отправить

Открыть LeaderBoard

ЗАГРУЗКА РЕШЕНИЙ НА ПЛАТФОРМУ

30 мая до 11:00

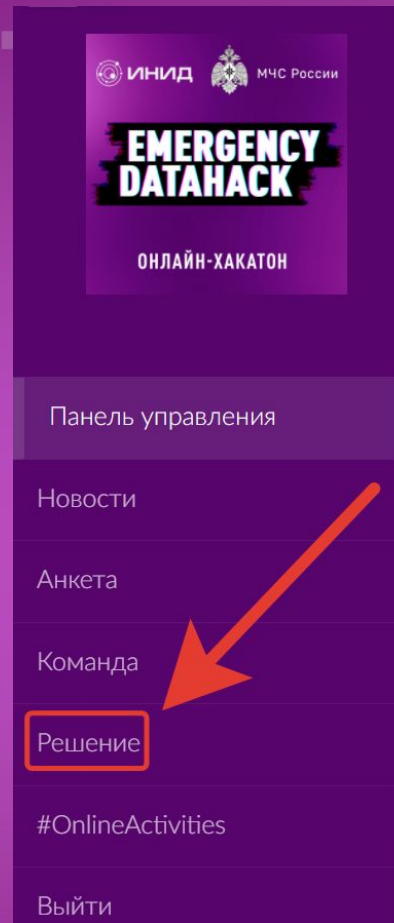
загрузить результаты на платформу во вкладку “Решения”

На защиту проектов будут допущены не более 10 команд согласно результатам лидерборда

Время защиты проекта:

5 минут - презентация проекта

3 минуты - ответы на вопросы



ПРОГРАММА ХАКАТОНА

28 мая

18:00 - Открытие хакатона

18:20 - Выдача презентаций с задачами и данных

19:30 - Q&A сессия с экспертами по задачам

29 мая

11:00 - Чек-поинт #1

15:00 - Мастер-класс от компании Tele2 “Геоаналитика на транспорте”

18:00 - Чек-поинт #2

30 мая

11:00 - Дедлайн загрузки решений

11:00 - 15:00 - Проверка кода

15:00 - 17:00 - Защита проектов

17:00 - 18:00 - Подведение итогов

18:00 - Объявление победителей, завершение хакатона

ОНЛАЙН-ХАКАТОН

EMERGENCY DATA HACK

УСПЕХОВ



МЧС
РОССИИ



ИНИД

