

BEST НАСХ'23

ТРЕБОВАНИЯ К РЕШЕНИЮ:

Перед отправкой решения в архиве (**название_команды.zip**) убедитесь, что оно соответствует следующим критериям:

1. Присутствует презентация решения в формате – **название_команды.pptx**, в которой присутствует следующая информация:

- Название команды
- Имя капитана команды
- ФИО членов команды
- Подробное текстовое описание решения с обоснованием выполненных действий
- Визуализация данных и работы модели
- Результаты, полученные при вычислении указанной метрики качества.

2. Присутствует файл с прогнозом для тестовых данных в формате – **название_команды.csv**

Решение отправляйте на почту alik.nigmatullin@bmstu-best.ru с темой [ПФ][Data Science] <Название команды>. В теле письма: обязательно должны быть указаны ФИО капитана команды.

**Все решения принимаются
до 27 апреля, 23:59.**

BEST НАСН'23

ЗАДАЧА

В ЖД логистике один из главных компонентов успешного функционирования бизнеса – это качество вагонного парка. Ведь именно хорошее состояние вагонов позволяет предоставлять клиенту безупречный сервис и выполнять договорные обязательства. А для поддержки парка рабочем состоянии важно своевременно и качественно проводить его ремонт.

До внедрения цифровых инструментов этот процесс был довольно консервативным, а большая его часть происходила в полуавтоматическом режиме со значительной долей ручных корректировок. Всё это вело к таким рискам, как выбор отдаленного депо (а это рост стоимости перегона), отсутствие необходимых деталей, непроизводительный простой в их ожидании и тд. И эти риски, в свою очередь, вели к дополнительным затратам.

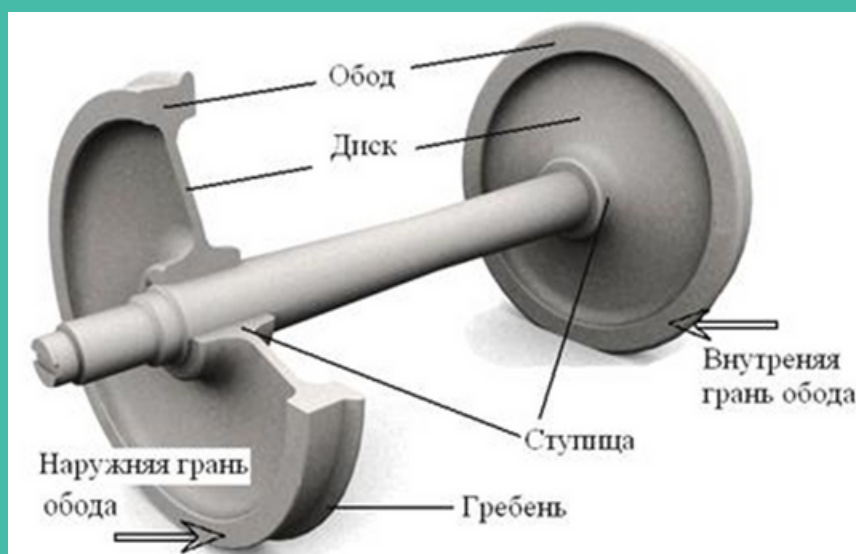
Не так давно ЖД отрасль начала двигаться в сторону цифровизации и сбора информации о состоянии вагона. На ЖД сети начали массово устанавливать датчики, которые измеряют технические показатели колёсных пар вагонов – одну из самых дорогостоящих его частей – и позволяют видеть информацию практически в режиме «онлайн».

BEST НАСЧЁТ

ЗАДАЧА (ПРОДОЛЖЕНИЕ)

ПГК Диджитал активно занимается внедрением предиктивной аналитики для контроля и управления своим вагонным парком. Мы разработали ряд предиктивных и оптимизационных алгоритмов, которые помогают нашим коллегам из Диспетчерского центра своевременно заадресовывать вагоны в ремонт, предупреждать о наличии серьёзных дефектов, планировать оптимальные маршруты в вагонно-ремонтные предприятия.

Наша компания активно собирает данные с датчиков на ЖД сети о техническом состоянии колёсных пар в режиме реального времени, тем самым собирая историю измерений о каждом нашем вагоне. Для своевременной забраковки вагона в ремонт, мы предлагаем вам решить задачу прогнозирования остаточного ресурса колёсных пар вагона т.е. **через сколько км** толщина гребня колёсной пары достигнет критического значения, используя данные контрольно-технических измерений.



BEST НАСХ'23

ЗАДАНИЕ

Необходимо разработать модель машинного обучения, способную прогнозировать остаточный ресурс колёсной пары на заданной тестовой выборке.

СЛОЖНОСТИ, С КОТОРЫМИ ПРИДЕТСЯ СТОЛКНУТЬСЯ:

- Показания толщины гребня колёсной пары имеют выбросы из-за погрешности измерительных приборов.
- Сбор данных с датчиков происходит с разной периодичностью (гранулярностью), поэтому могут быть «пробелы» в данных.
- Присутствуют пропуски в данных из-за особенностей считывания и измерения технических показателей.
- Небольшая глубина данных для обучения модели. Мы должны заранее предсказывать остаточный ресурс вагона, поэтому доступны только первые 19 показаний о толщине гребня колёсной пары.

BEST НАСН'23

КРИТЕРИЙ ОЦЕНКИ:

Качество модели будет оцениваться по метрике MAE (Mean Absolute Error).

ДААННЫЕ К ЗАДАЧЕ:

Датасет к задаче и дополнительную информацию вы можете получить по ссылке:

<https://drive.google.com/file/d/1BkM5fNS1yVA1QcAK62TgSQ1nWUWjbg3b/view>

ДОПОЛНИТЕЛЬНЫЕ БАЛЛЫ:

Дополнительные баллы можно получить, если алгоритм будет иметь минимальный web-интерфейс, чтобы пользователь при загрузке файла с техническими измерениями мог получить предсказание алгоритма – остаточный ресурс вагона.