



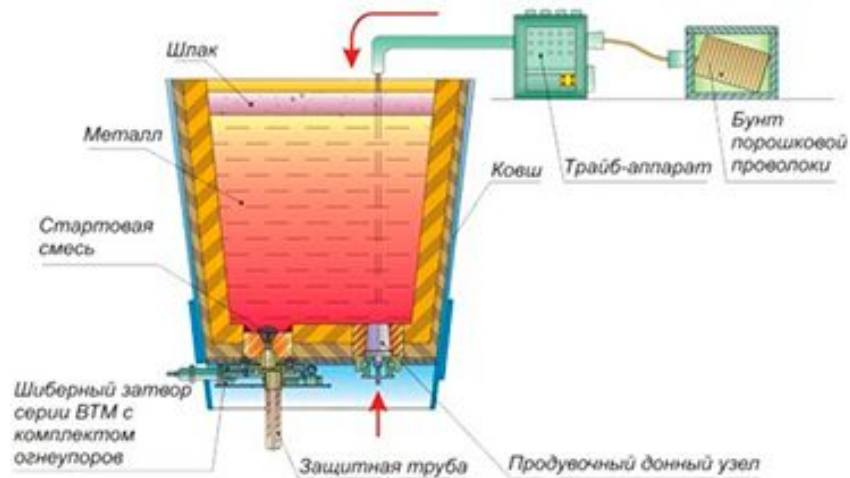
Оптимизация процесса легирования сталей

Куратор:
Станислав Васильев, ЕВРАЗ-Техника

Описание:
Изучение реальных данных физико-химического процесса легирования сталей. Создание алгоритма определения состава шлака или алгоритма оптимизации расхода извести.

Авторы:
Дмитрий Городецкий, Евгений Заварзин, Кирилл Келлер, Александр Мелехин, Анастасия Савкина, Анвар Хафизов

Введение



<https://vulkan-tm.com/tehnologii/obrabotka-metalla-argonom-v-kovshe/>

“Мы варим шлак”

“Чтобы сварить хорошую сталь — нужно сварить хороший шлак”

Цели проекта

1. Провести исследовательский анализ данных
2. Разработать ML-модель, предсказывающую химический состав последней пробы шлака
3. Проанализировать возможность оптимизации расхода извести с помощью методов ML

Цели проекта

1. Провести исследовательский анализ данных
2. Разработать ML-модель, предсказывающую химический состав последней пробы шлака
3. Проанализировать возможность оптимизации расхода извести с помощью методов ML

Цели проекта

1. Провести исследовательский анализ данных
2. Разработать ML-модель, предсказывающую химический состав последней пробы шлака
3. Проанализировать возможность оптимизации расхода извести с помощью методов ML

Работа в команде



Google
Sheets



Этап анализа данных

1. Изучение формы распределения данных
2. Удаление пропусков
 - а. Объекты и признаки с большим количеством пропусков
 - б. Объекты с пропусками в целевых признаках
3. Удаление выбросов
4. Кодирование категориальных признаков и заполнение пропусков

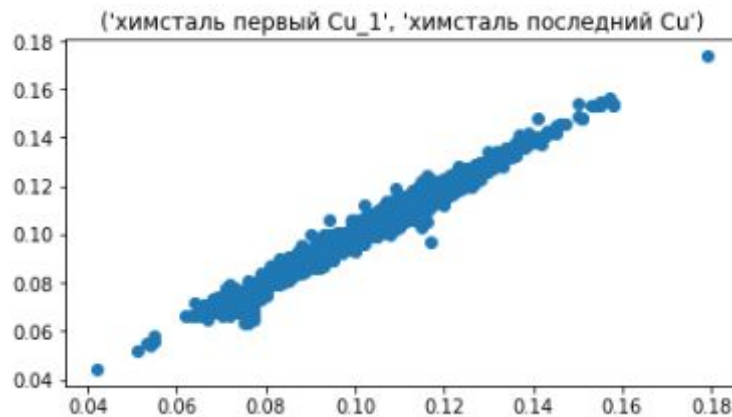
Результат предобработки:

(7041, 84) -> (3003, 66) – отбросили 50+% объектов, 18 признаков

Предсказание состава шлака

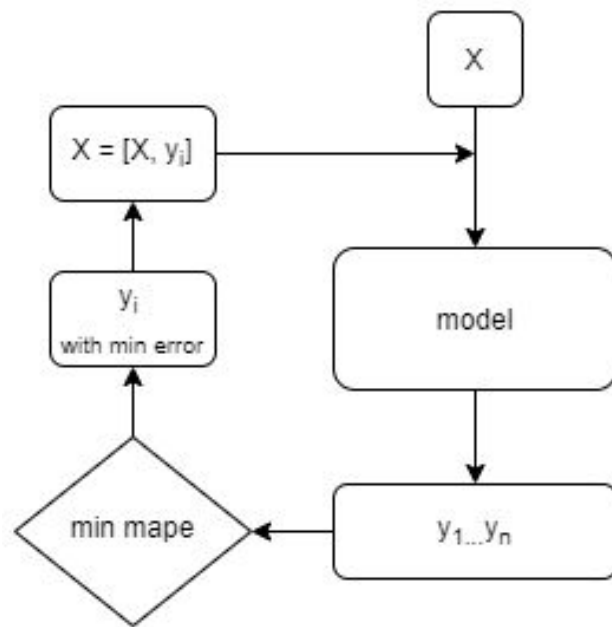
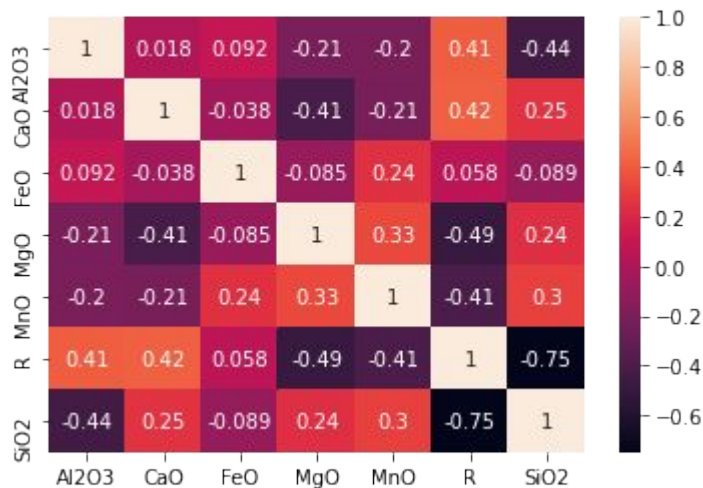
Общий подход

- удаление линейно зависимых столбцов
- удаление выбросов методом кластеризации
- удаление признаков через backward elimination
- выбор модели (LinReg, GB, LGBM, RandomForest)



Итеративный подход

Корреляция между целевыми переменными



Оптимизация расхода извести

Новый целевой признак - объем подаваемой сыпучей извести

Модели классификации:

- RandomForest
- DecisionTrees
- CatBoost

Использованные инструменты



Итоговый результат

Предсказание состава шлака

Цел. призн.	химшлак последний Al ₂ O ₃	химшлак последний CaO	химшлак последний FeO	химшлак последний MgO	химшлак последний MnO	химшлак последний R	химшлак последний SiO ₂
MAPE _{общий}	9,9%	4,7%	28,5%	19,2%	27%	5,6%	6,6%
MAPE _{итер}	2,68%	0,75%	6,50%	5,51%	7,15%	0,40%	0,76%

Итоговый результат

Оптимизация расхода извести

Выбор диапазона расхода извести по целевой химии

Модель	RandomForest	DecisionTrees	CatBoost
accuracy	47%	47%	56%
balanced accuracy		-	34%

Задача нуждается в дальнейшем изучении...

Заключение

- НАУЧИЛИ МОДЕЛИ ПРЕДСКАЗЫВАТЬ ШЛАК
- Не надо восстанавливать таргеты регрессией
- Можно думать, что ошибся, но не ошибиться



Контакты

Савкина Анастасия



@as_savkina



savkina.anastasia.451@gmail.com

Мелехин Александр



@metra4ok



amelekhin96@gmail.com

Хафизов Анвар



@ankhafizov



Khafizov.A@crys.ras.ru

Кирилл Келлер



@keller_kirill



kellertab@gmail.com

Заварзин Евгений



@Evgenius2020



zavarzinevg@gmail.com

Городецкий Дмитрий



@aka_naked_gun



gorodeckiydimchik@gmail.com