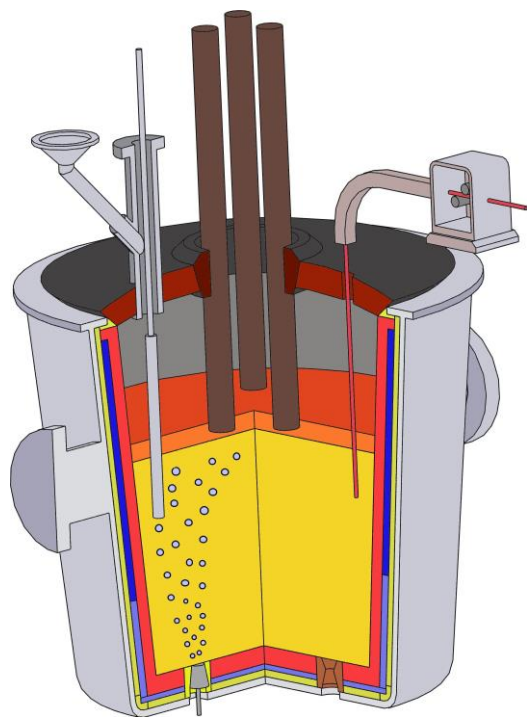




Отборочное задание

«Легирование сталей. Прогнозирование химического состава шлака»

март 2021 г



Сталь обрабатывают в металлическом ковше вместимостью около 110 тонн. Чтобы ковш выдерживал высокие температуры, изнутри его облицовывают огнеупорным кирпичом. Расплавленную сталь заливают в ковш и подогревают до нужной температуры графитовыми электродами. Они установлены в крышке ковша.

Из сплава выводится сера (десульфурация), добавлением примесей корректируется химический состав и отбираются пробы. Сталь легируют — изменяют её состав — подавая куски сплава из бункера для сыпучих материалов или проволоку через специальный трайб-аппарат (англ. tribe, «масса»).

Перед тем как первый раз ввести легирующие добавки, измеряют температуру стали и производят её химический анализ. Потом температуру на несколько минут повышают, добавляют легирующие материалы и продувают сплав инертным газом. Затем его перемешивают и снова проводят измерения. Такой цикл повторяется до достижения целевого химического состава и оптимальной температуры плавки.

Тогда расплавленная сталь отправляется на доводку металла или поступает в машину непрерывной разливки стали МНЛЗ. Оттуда готовый продукт выходит в виде заготовок-рельс, каждая из которых пилится на 4 отрезка - крата.

Задача для отборочного этапа

Изучите данные физико-химического процесса легирования сталей. Создайте алгоритм определения химического состава шлака по исходным данным.

Данные

Предоставлены реальные агрегированные данные по химии и физике процесса. Файл «Исходные данные.csv» (данные по плавкам о количестве добавок, расходе э/э, химическом анализе стали и хим анализ шлака, забираемых в начале плавки и в конце). Данные требуется очистить, в трети выборки показатели хим анализа отсутствуют, их нужно занулить. Зафиксировать самую популярную по плавкам марку стали и для нее спрогнозировать целевой состав шлака.

План работ

EDA – выводы по качеству данных:

1. Пропуски
2. Нули
3. Дисперсия
4. Выбросы

Провести EDA (исследовательский анализ данных) целевого химического состава шлака и предсказать:

- химшлак последний Al_2O_3
- химшлак последний CaO
- химшлак последний R
- химшлак последний SiO_2

Необходимые компетенции участников

Навыки по очистке и аугментации данных

Навыки построения статистических или эвристических моделей

Понимание задачи регрессии и её метрик

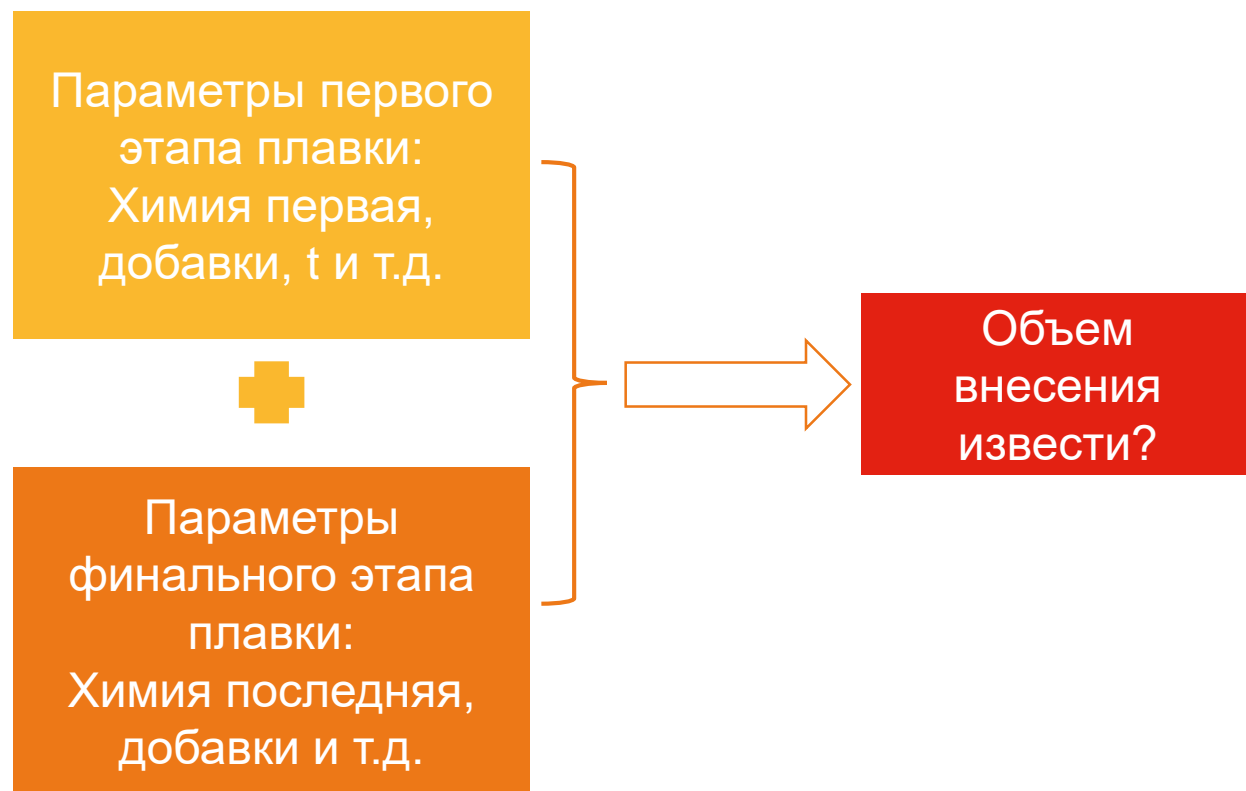
Рекомендуется знание языка программирования Python и пакетов pandas, scikit-learn

Для самой популярной марки стали и для фиксированной химии шлака (название столбца «химшлак_последний») необходимо предсказать объем внесения извести 'сыпуч_известь_рп' для получения этой целевой (фиксированной) химии шлака.

Данные

Все данные из датасета (без нового таргета). Таргетом является 'сыпуч_известь_рп'.

Предсказание отдачи извести



Метрики качества для финального задания

Модель	R2	MAE	MSE	MAPE	Лучшие параметры модели
Linreg...					
Gradientboosting					
.....					

Коротко напишите - какой метрикой по вашему мнению следует руководствоваться при выборе модели в первую очередь и почему?

Вывод – какая модель в вашей табличке лучше и почему?