

Кейс компании Garrix Результаты.

Команда 6:

Бобров Александр

Бобров Максим

Богомольный Александр

Каприелов Юрий

Макаров Александр

Шкурихин Михаил



«Дорожная карта работ по кейсу.»

1. Осмыслена задача и произведен анализ входных данных;
2. Составлен план работ и подготовлена дорожная карта, распределены роли в команде;
3. Данные спарсены и переведены в табличный формат для дальнейшей работы;
4. Проанализированы полученные данные, найдены корреляционные зависимости.
5. Опробованы два решения - XGBoost и нейронная сеть.

«Решение»

Наша гипотеза заключается в том, что на укладку коробок будет влиять их неоднородность по их геометрическим параметрам. И поскольку мы не знаем как коробки будут располагаться в итоговой укладке мы выделили следующие параметры: объем, вытянутость (отношение минимальной площади к оставшейся стороне), «плоскость» (отношение максимальной площади к стороне). Таким образом мы избавились от необходимости подбирать укладку.

Следующим шагом мы рассчитали различные статистические параметры, как внутри набора коробок (средние значения параметров, медианы, дисперсию), так и по отношению к общему объему и объему контейнера.

Выделили еще один тип коробок, максимальная сторона которых превышает минимальную сторону контейнера (что ограничивает нас в способе укладки)

Отдельно для специфических коробок (превышающих минимальный размер, не штабелируемых и не переворачиваемых) посчитали отношение их суммарных объемов к общему объему, объему контейнера, количества таких коробок в наборе.

Команда разработала два решения **XGBoost** и **нейронную сеть**.



«Параметры»

Параметр	Код	Название в коде	Условный тип
Ширина контейнера	f0	cargo_width	Контейнер
Длина контейнера	f1	cargo_height	Контейнер
Высота контейнера	f2	cargo_length	Контейнер
Площадь контейнера	f3	cargo_area	Контейнер
Объем контейнера	f4	cargo_volume	Контейнер
Суммарный объем коробок	f5	boxes_volume	Объем
Средний объем коробок в наборе	f6	boxes_volume_mean	Неоднородность объемов
Медианный набор коробок в наборе	f7	boxes_volume_median	Неоднородность объемов
Дисперсия распределения коробок	f8	boxes_volume_var	Неоднородность объемов
Отношение среднего к медианному	f9	boxes_volume_lag	Неоднородность объемов
Средневзвешенное среднее объемов коробок в общем объеме груза	f10	boxes_volume_norm_vs_boxes_mean	Объем
Средневзвешенное среднее объемов коробок в общем объеме контейнера	f11	boxes_volume_norm_vs_cargo_mean	Объем
Доля нештабилируемых коробок от объема контейнера	f12	stacking_volume_cargo_perc	Нештабилируемые коробки
Доля нештабилируемых коробок от общего объема груза	f13	stacking_volume_boxes_perc	Нештабилируемые коробки
Доля коробок размеры которых превышают размеры контейнера	f14	oversize_volume_boxes_perc	Оверсайз коробки
Доля непереворачиваемых коробок от объема контейнера	f15	turnover_volume_cargo_perc	Непереворачиваемые коробки
Доля непереворачиваемых коробок от общего объема груза	f16	turnover_volume_boxes_perc	Непереворачиваемые коробки



«Параметры»

Параметр	Код	Название в коде	Условный тип
Доля площади непереворачиваемых коробок от общего объема контейнера	f17	turnover_area	Неперворачиваемые коробки
Отношение максимальной высоты непереворачиваемой коробки к высоте контейнера	f18	turnover_height	Неперворачиваемые коробки
Максимальная плоскость	f19	flatness_max	Неоднородность по линейным параметрам
Средняя плоскость	f20	flatness_mean	Неоднородность по линейным параметрам
Медианная плоскость	f21	flatness_median	Неоднородность по линейным параметрам
Дисперсия плоскости	f22	flatness_var	Неоднородность по линейным параметрам
Отношение средней плоскости к медиане	f23	flatness_lag	Неоднородность по линейным параметрам
Максимальная вытянутость	f24	lengthness_min	Неоднородность по линейным параметрам
Средняя вытянутость	f25	lengthness_mean	Неоднородность по линейным параметрам
Медианная вытянутость	f26	lengthness_median	Неоднородность по линейным параметрам
Дисперсия вытянутость	f27	lengthness_var	Неоднородность по линейным параметрам
Отношение средней плоскости к медиане	f28	lengthness_lag	Неоднородность по линейным параметрам
Доля нештабилируемых коробок	f29	mean_stacking	Количество
Доля непереворачиваемых коробок	f30	mean_turnover	Количество
Количество коробок в наборе	f31	boxes_qty	Количество

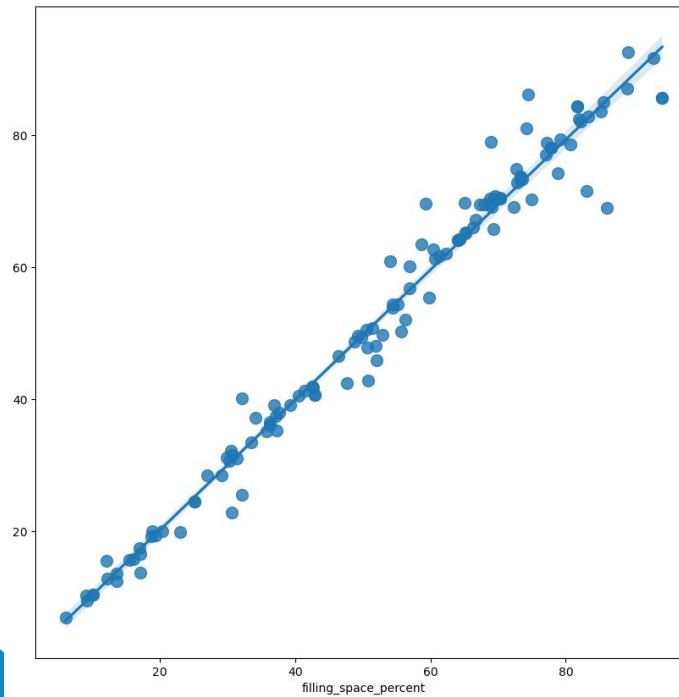


«Решение»

Результаты:

XGBoost

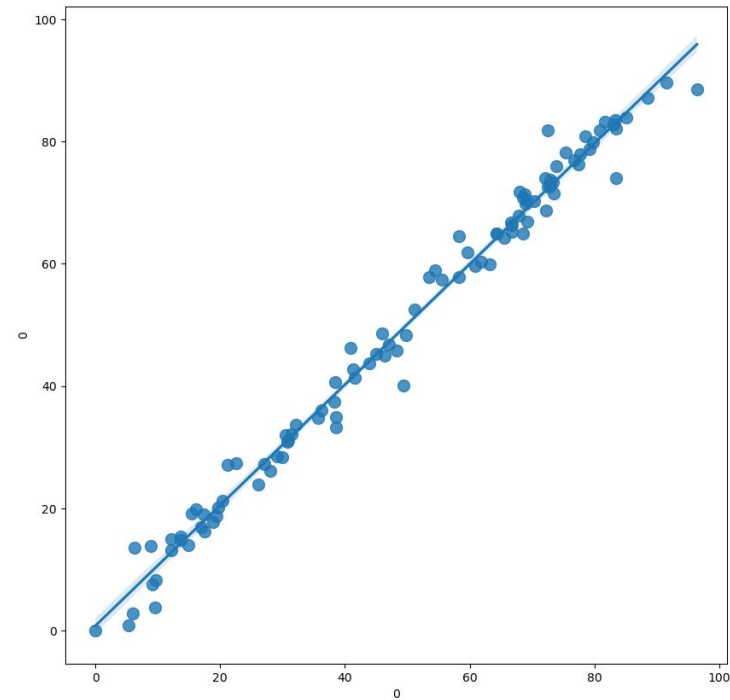
XGBoost $R^2=0.9754$



Нейронная сеть

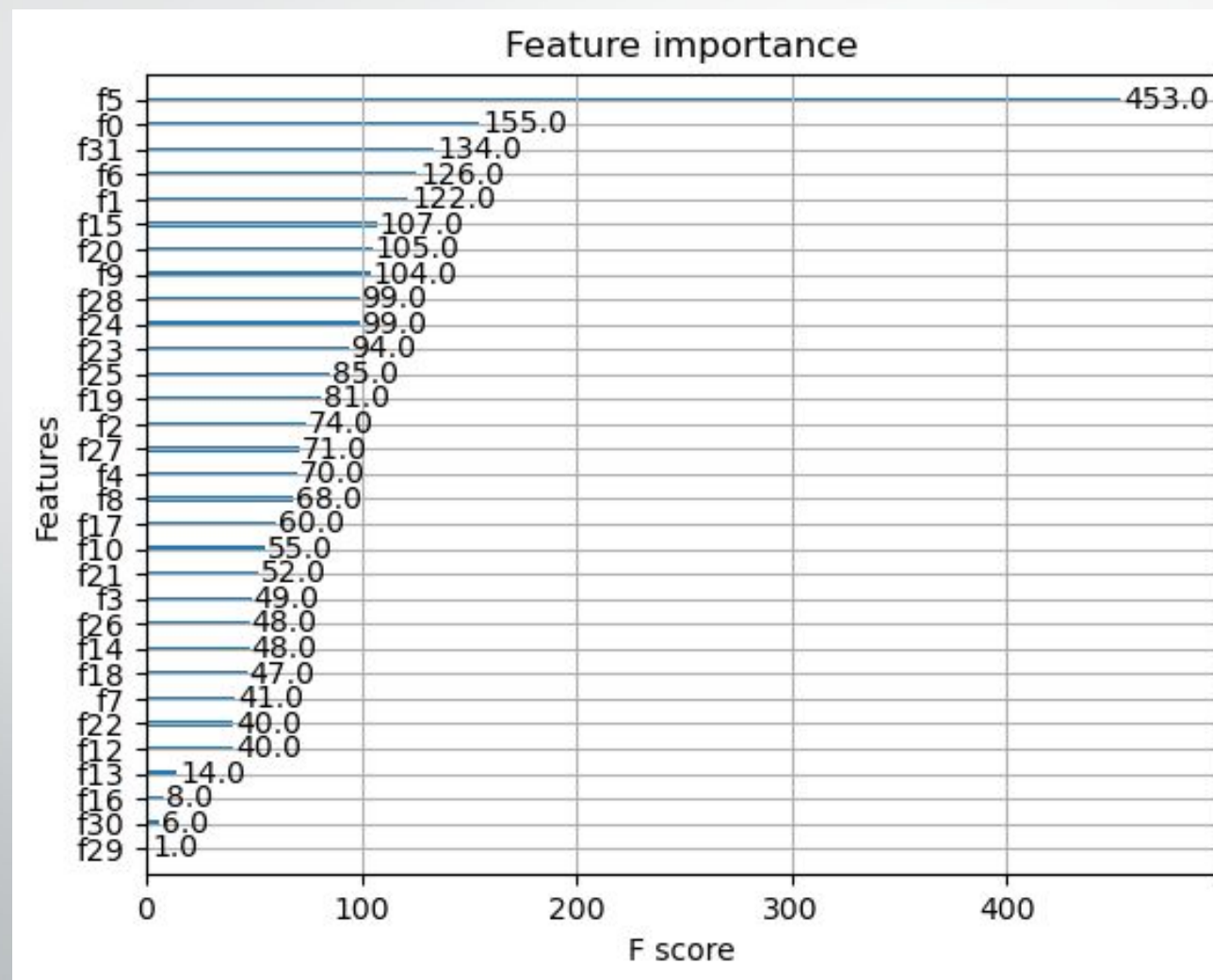
2 внутренних слоя по 500 нейронов

XGBoost $R^2=0.9863$ MSE:8.9402



«Решение»

Вес параметров в модели XGBoost



«Преимущества»

- а) Повышение эффективности: прогнозируется максимально возможное значение целевого показателя тем самым позволяя максимально загружать контейнеры и перевозить больше.
- б) Сокращение отходов. Больше перевозим товара за один раз - меньше отходов и затрат на топливо и т.д.
- с) Повышение конкурентоспособности: за счет снижения затрат улучшаем производительность труда, тем самым можем сдерживать цены и зарабатывать на объеме.



«Репозитории GitHub»

Нейронная сеть (train/test):

https://github.com/fahtom94/garpix-hackathon/blob/main/Result/NN_train_test.ipynb

Нейронная сеть (на всем множестве):

https://github.com/fahtom94/garpix-hackathon/blob/main/Result/NN_all.ipynb

Веса нейронной сети (на всем множестве)

https://github.com/fahtom94/garpix-hackathon/blob/main/Result/NN_model_weights.pth

XGBost модель:

https://github.com/fahtom94/garpix-hackathon/blob/main/Result/XGBoost_model.ipynb

XGBost параметры:

https://github.com/fahtom94/garpix-hackathon/blob/main/Result/xgboost_params.csv

Преобразование данных:

<https://github.com/fahtom94/garpix-hackathon/blob/main/Result/Pars%2BFeatures.ipynb>

Описание «внутренних» метрик

https://github.com/fahtom94/garpix-hackathon/blob/main/Result/Parameters_description.xlsx

