



DS-prof-2022-08

Меня хорошо видно && слышно?



Защита проекта

Тема: Прогнозирование вероятности наступления целевого события на трубопроводах ресурсоснабжающей организации



Дьяконов Семен

Начальник управления балансов и энергоресурсов
ООО «НОВОГОР-Прикамье» г. Пермь

План защиты



Цели проекта

The diagram shows a vertical list of five blue rounded rectangular boxes. A dashed line on the left side of the boxes connects them in a continuous path, with small arrowheads pointing downwards at each connection point, indicating a sequential flow from top to bottom.

Что планировалось

Используемые
технологии

Что получилось

Выводы

Цели проекта

1. Построение модели предсказания аварийных сетей для планирования ремонтов, перекладок и т.п. → снижение утечек и объема NRW (не проданная вода);
2. Построить «хорошую» модель. Применить модель классификации на реальных данных;
3. Изучить catboost

Что планировалось

1. В начале были Excel, VBA и все обработки данных был в них. Собрать данные из разные таблицы
2. Подключить Pandas и обработать данные
3. Подключить модели машинного обучения и посчитать вероятность
4. Проект делал 2 месяца. Первые 1.5 месяца выгружал данные из базы данных, изучал переводил из *.dbf в *.xlsx. Вторые 0,5 месяца делал проект

Используемые технологии

1. АИАС ОСВВ. База данных ООО «НОВОГОР-Прикамье». Anaconda jupyter notebook, Excel, *.dbf, *.xlsx
2. Библиотеки без "conda install": pathlib, Pandas. Numpy, matplotlib, seaborn, sklearn (LabelEncoder, train-test-split*, MinMaxScaler, RandomizedSearchCV), scipy
3. Библиотеки без "conda install" from sklearn.metrics import accuracy-score*, precision-score*, recall-score*, f1-score*, roc-auc-score*
4. Библиотеки с "conda install": Catboost. Phik
5. Catboost. Чудесно! Сократило создание модели, удалось взять выходной и поспать.

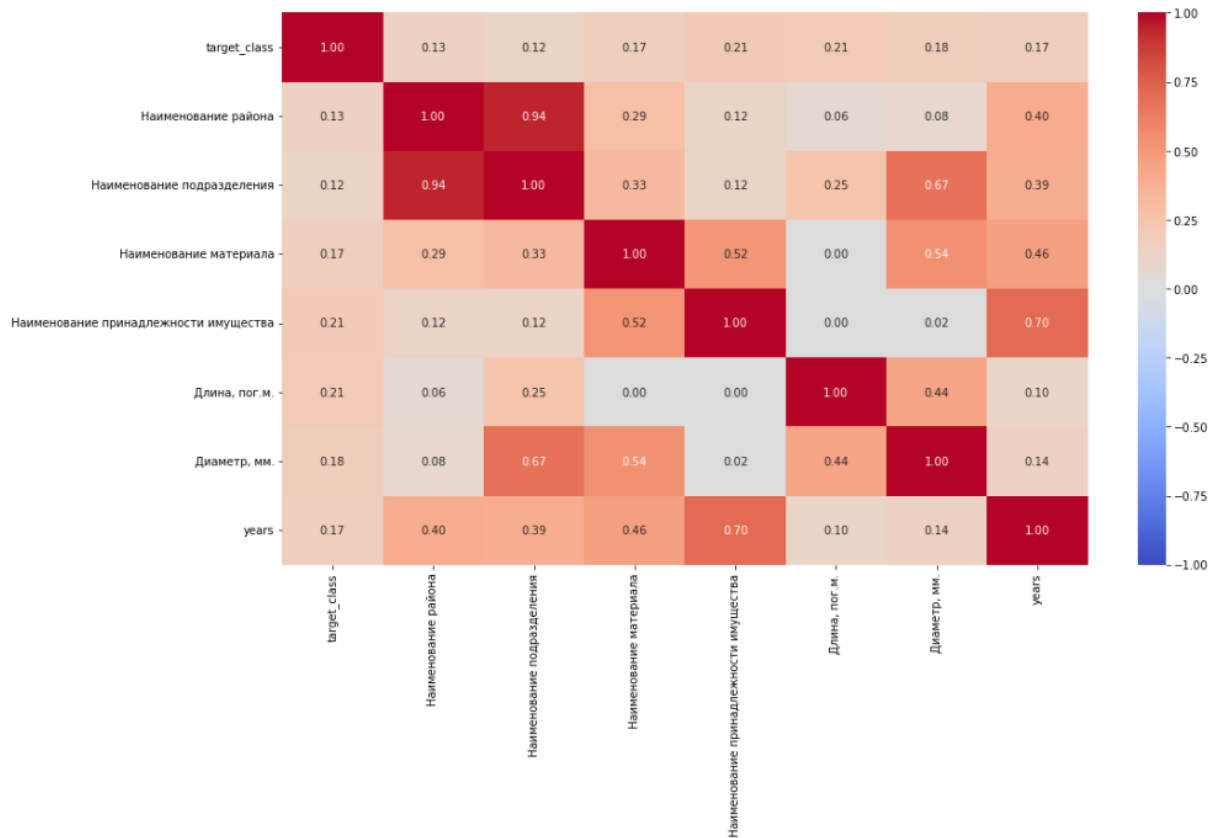
**минус заменить на подчеркик*

Что получилось

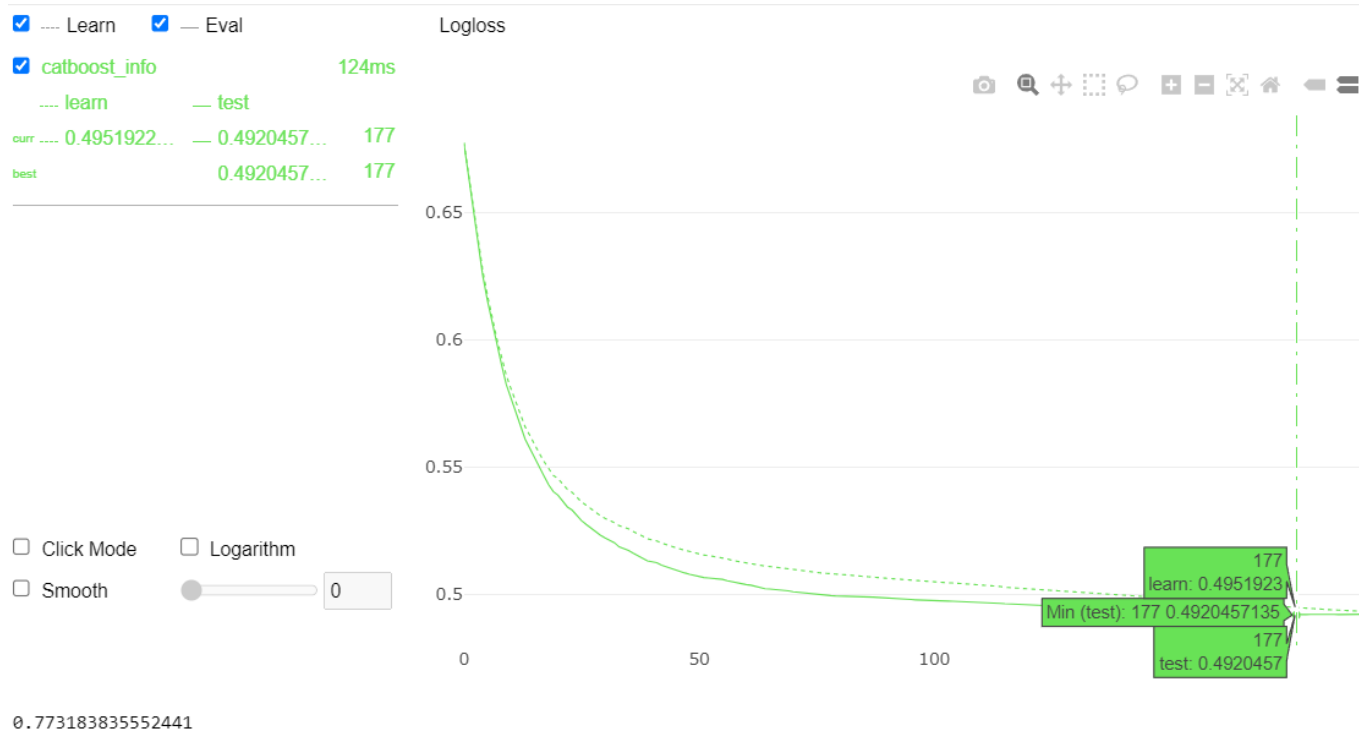
1. Ссылка на репозиторий с исходным кодом:

https://github.com/smddyakonov/otus_ds_dyakonov/tree/main/hw-7

Что получилось. Phik корреляции



Что получилось. График обучения



Выводы и планы по развитию

1. По ходу реализации пришлось поменять метод прогнозирования и отказаться от погодных влияний;
2. С учетом изменений п.1 достиг поставленной цели: спрогнозированы сети для формирования предложений на переключку в ближайший срок тарифного регулирования;
3. Расширить число влияющих переменных. Провести более глубокую проработку вопроса с учетом развития на работе центра автоматизированного управления

Спасибо за внимание!