



DS-prof-2022-08

Меня хорошо видно && слышно?



Защита проекта

Тема: Прогнозирование электропотребления предприятия с применением искусственных нейронных сетей

Здесь
могла
бы
быть
ваша
реклам
а

Здесь
могла
бы
быть
ваша
реклам
а

Дьяконов Семен

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

План защиты

Цели проекта

Что планировалось

Используемые
технологии

Что получилось

Выводы



**Здесь
могла бы
быть ваша
реклама**

Цели проекта

0. Актуальность обусловлена необходимостью исследование возможности применения методов искусственного интеллекта для краткосрочного прогнозирования электропотребления групп точек поставки электроэнергии предприятиям. С целью применения таких методов в помощь к экспертному прогнозированию.

1. Исследование применения искусственных нейронных сетей для краткосрочного прогнозирования потребления электроэнергии предприятием;

2. Изучить Рекуррентные нейронные сети (РНС, англ. Recurrent neural network, RNN), Длинные цепи элементов краткосрочной памяти (англ. Long short-term memory; LSTM)

3. Сравнить базовый прогноз и прогноз построенный с применением искусственных нейронных сетей;

Что планировалось



1. Сбор и анализ данных



2. Анализ временного ряда



3. Моделирование



4. Оценка результатов



5. Выводы

Используемые технологии

1. Файлы в формате *.xlsx
 2. Excel, python стандартные библиотеки, pytorch, RNN, LSTM
-

Что получилось. Репозиторий на github.

1. Ссылка на репозиторий с исходным кодом:

https://github.com/smddyakonov/otus_ds_dyakonov/tree/main/hw-19

Что получилось. Описание исходных данных

Почасовое потребление электрической энергии представлено

- за период с 2018 по июнь 2023 г.;
- в разрезе точек поставки;
- в разрезе дата-время

Предварительные преобразования для моделирования

Сделано в отдельном ноутбуке

- сформировать список исходных файлов
- проверить, что во всех исходных файлах есть лист "Данные эл.эн.". В листе "Данные эл.эн." хранятся почасовые расходы электроэнергии по точкам поставки
- объединить точки поставки (ТП) по годам и найти перечень всех ТП за период с 2019 - 2023 гг
- добавить привязки ТП - Объект - группа точек поставки (ГТП)

Что получилось.

Ввод [7]:

1 df_ee

Out[7]:

	Дата время (мск)	ГТП_ВС	ГТП_ВО
0	2019-01-01 00:00:00	6344.080	8562.00
1	2019-01-01 01:00:00	6311.008	8170.24
2	2019-01-01 02:00:00	6345.960	7837.36
3	2019-01-01 03:00:00	6190.032	7835.24
4	2019-01-01 04:00:00	6354.032	7821.12
...
39379	2023-06-29 19:00:00	7132.000	6772.00
39380	2023-06-29 20:00:00	7058.000	6860.00
39381	2023-06-29 21:00:00	6967.000	7041.00
39382	2023-06-29 22:00:00	6265.000	6772.00
39383	2023-06-29 23:00:00	5696.000	6664.00

39384 rows × 3 columns

В итоге получил дата-фрейм:

Колонка «Дата время (мск)» – дата-время, часовой пояс – «МСК»

Колонка «ГТП_ВС» - потребление электрической энергии за каждый час, по группе точек поставок (ГТП) водоснабжение (ВС);

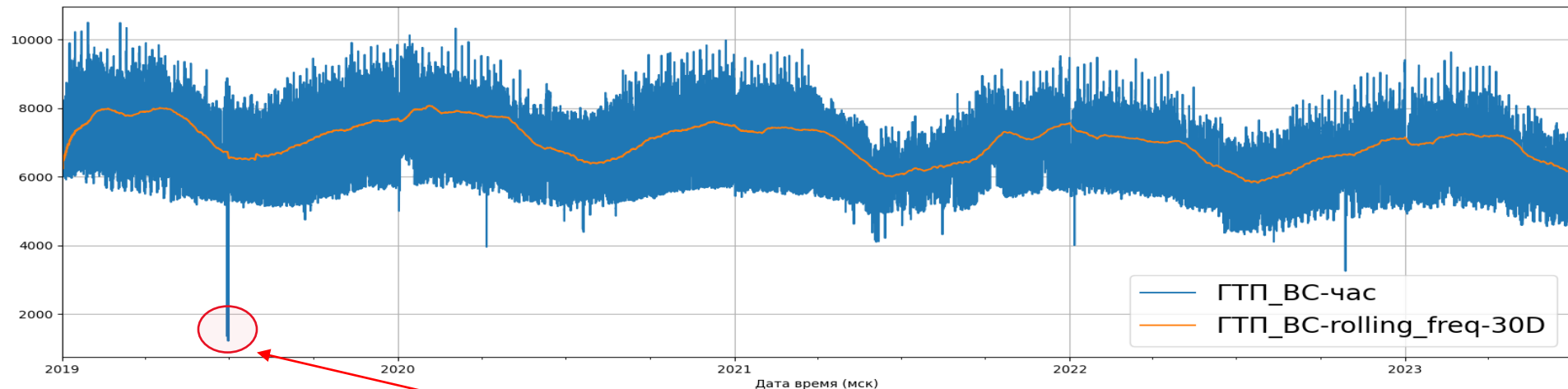
Колонка «ГТП_ВО» - потребление электрической энергии за каждый час, по группе точек поставок (ГТП) водоотведение (ВО);

Всего строк: 39 384

Тип данных: временной ряд с частотой 1 час



Что получилось. Предобработка данных (Data Preprocessing)

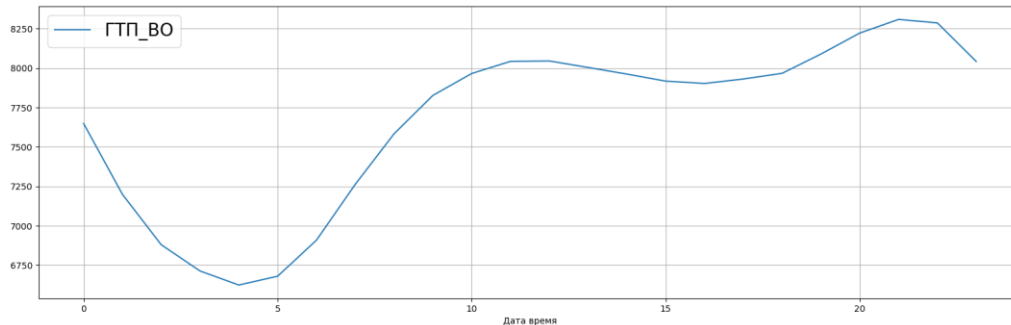
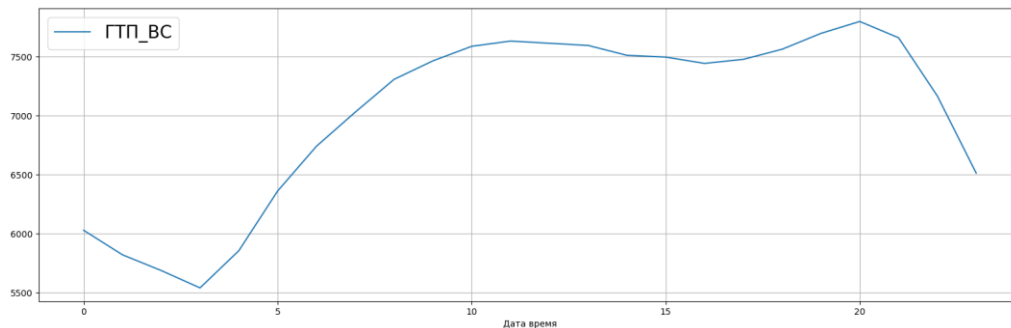


**Проверка на выбросы, обнаружено не характерно поведение данных. Технологический останов сооружений. Замену на среднее значение аналогичного года, месяца, часа.
По ГТП_ВО выброс не обнаружено**

Что получилось. EDA

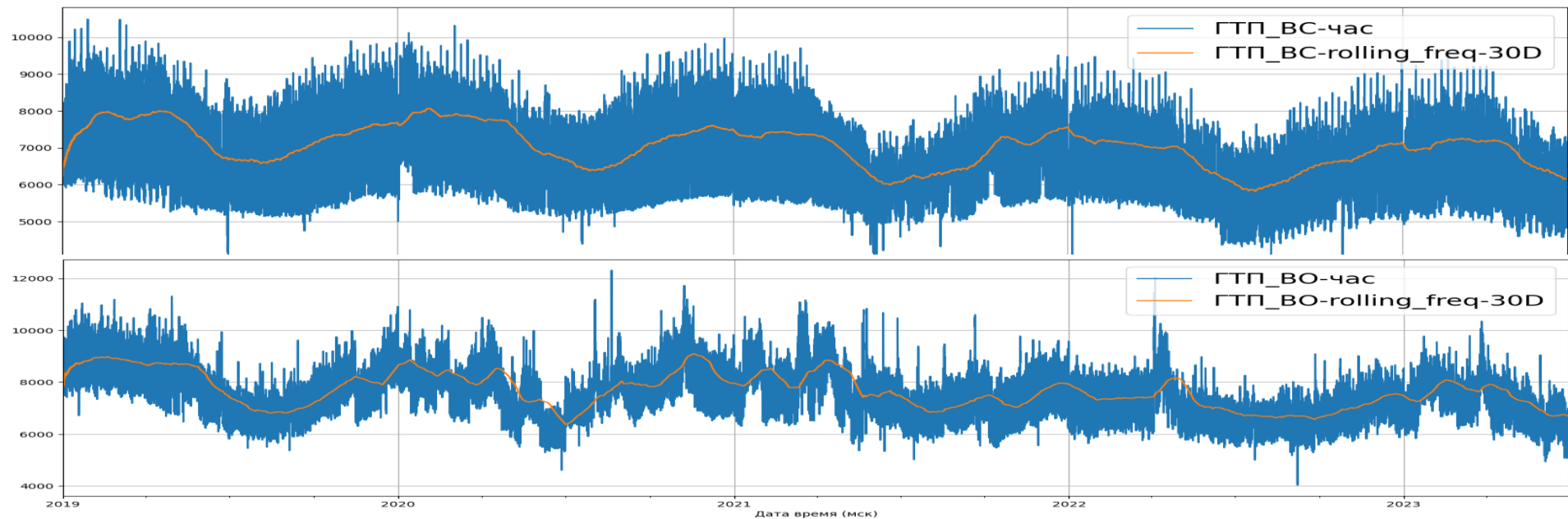
```
1 df_ee_eda.describe()
```

	ГТП_ВС	ГТП_ВО
count	39384.000000	39384.000000
mean	7025.898740	7667.255423
std	1067.284958	989.238861
min	4000.000000	4030.000000
25%	6208.000000	6969.000000
50%	7004.000000	7612.000000
75%	7844.898000	8291.000000
max	10493.112000	12312.840000



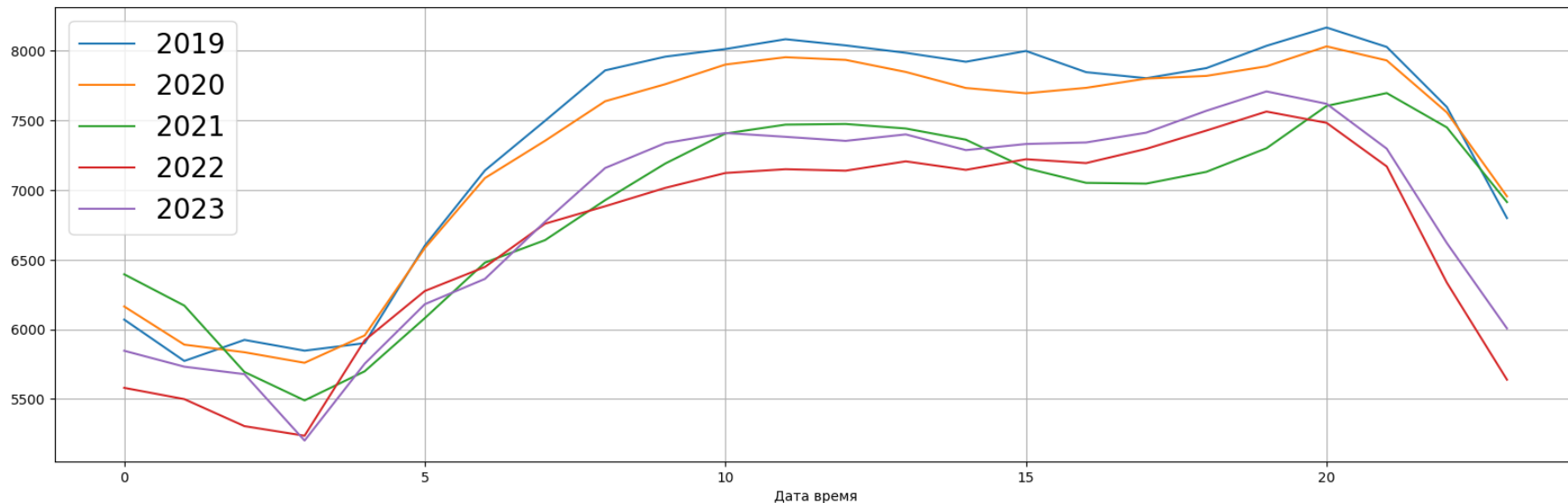
Что получилось. «Чистые данные»

Профиль потребления электроэнергии ГТП_ВС, ГТП_ВО



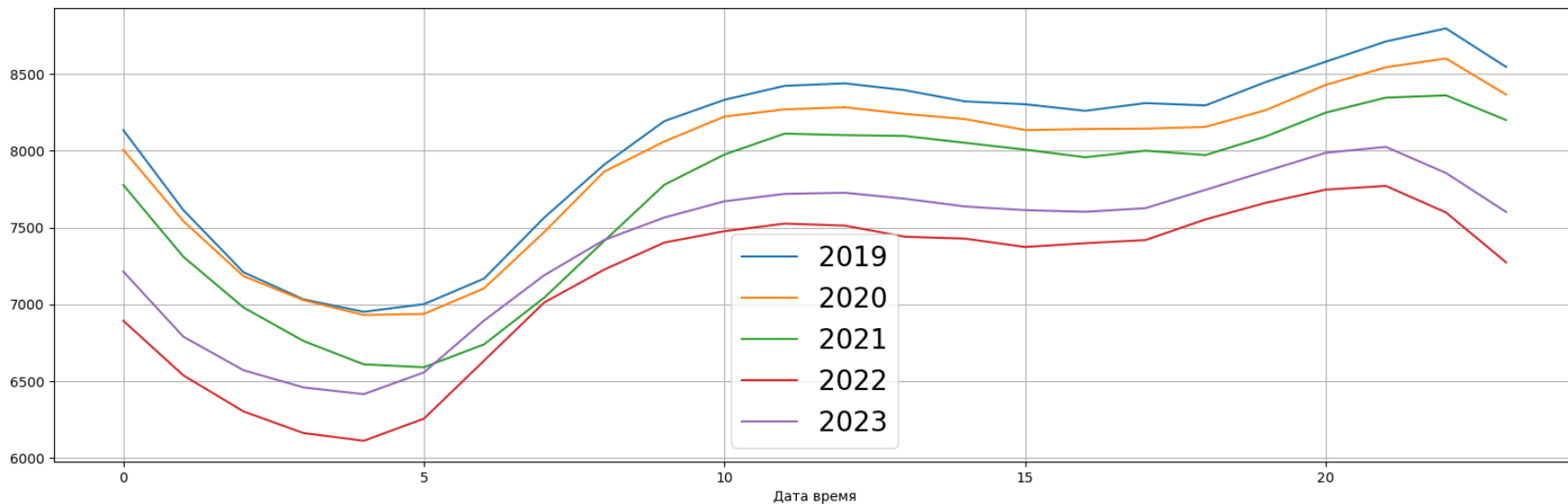
Что получилось. «Чистые данные» ГТП_ВС

Усредненный профиль потребления электроэнергии ГТП_ВС



Что получилось. «Чистые данные» ГТП_ВС

Усредненный профиль потребления электроэнергии ГТП_ВС



Что получилось. Моделирование

Исследование применения искусственных нейронных сетей для краткосрочного прогнозирования потребления электроэнергии предприятием;

Прогнозирую сразу одни сутки, а в сутках 24 значения, то за основу возьму RNN интервальное прогнозирование

Базовая модель прогноз на сутки вперед равен аналогичному дню неделю назад

Гипотеза: прогнозирование с помощью ИИ должно быть лучше, чем базовая модель

Оценка подхода проводится по бизнес-метрике, БМ:

$$БМ = \frac{\sum_h^N \frac{\Phi - \text{Откл_abs}}{\Pi}}{N}$$

Φ - фактический расход электрической энергии за час; Π - плановый расход электрической энергии за час; Откл_abs = |Φ - Π| отклонение фактического расхода за час от планового расхода за час N - число плановых значений за месяц.

Выводы и планы по развитию

1. Цели достигнуты
2. Для прогнозирования почасового потребления электрической энергии можно использовать искусственные нейросети
Для ГТП_ВС на тесте ИИ-модель показала себя лучше в зимние и летние месяцы.
3. Применение трансформеров для задачи

Спасибо за внимание!