

Инструкция для проверки скриптов

Скопировать все файлы в одну папку для последующего запуска в командной строке.

1. Подготовка синтетических датасетов часового потребления для множественных ИТП (индивидуальных тепловых пунктов) – холодная вода, но то же самое и для горячей воды. Минимальные данные – дата, час, расход в м3, дополнительно - тип потребителя, сегмент и т.д. Датасет записывается в соответствующую поддиректорию. Запуск: `python generate_many_itps_with_types.py`
2. Подготовка данных для добавления в датасет (день недели, неделя года, лунный день, выходные/рабочие/праздничные, возможно добавление прочих параметров – среднесуточная температура, осадки, социальная и экономическая информация, общая для города/сегмента. Запуск: `python make_fixed_params.py`
3. Подготовка общего датасета и запуск обучения модели (CNN+Transformer encoder+FFNN) на датасете. Модель небольшая, поэтому обучение очень быстрое. Последняя и лучшая модели записываются в поддиректорию. Относительно медленная операция – подготовка тензоров. Требуется оптимизация или загрузка ранее сформированного датасета, код вариативен. Запуск: `python train_transformer_holiday.py`
4. `parquet_to_batches.py` – модуль создания батчей из датасета в формате Parquet
5. Выдача предсказаний потребления и визуализация на демонстрационном синтетическом датасете – модуль `demo_predict_plot.py`. Запуск: `python demo_predict_plot.py`