

Описание кода и файлов в папке airpollpredictor2

В папке airpollpredictor2 сохранен код написанный Шамсутдиновой Аделиной (<https://github.com/Adele-Kim>), а также промежуточные датасеты.

Описание выполненных задач:

- 1) Анализ метаданных https://github.com/AlexandraNasonova/air_pollution_predict/blob/master/airpollpredictor2/metadata_analysis.ipynb
- 2) Создание новых более полных метаданных https://github.com/AlexandraNasonova/air_pollution_predict/blob/master/airpollpredictor2/metadata_creator.ipynb
- 3) Написание скрипта для создания датасетов по выбранным странам/городам https://github.com/AlexandraNasonova/air_pollution_predict/blob/master/airpollpredictor2/dataset_creator.ipynb
- 4) Исследовательский анализ данных Дании и Сербии https://github.com/AlexandraNasonova/air_pollution_predict/blob/master/airpollpredictor2/eda_RS_DK.ipynb
- 5) Написание скрипта для экспорта информации о станциях в kml-формате для последующей визуализации станций с использованием Google Earth https://github.com/AlexandraNasonova/air_pollution_predict/blob/master/airpollpredictor2/kml_exporter.ipynb
- 6) Калькулятор AQI и код для создания итогового timeseries с AQI усредненным по дням https://github.com/AlexandraNasonova/air_pollution_predict/blob/master/airpollpredictor2/aqi_calculator_v1.ipynb
- 7) ML-эксперименты на данных одной выбранной станции Дании https://github.com/AlexandraNasonova/air_pollution_predict/blob/master/airpollpredictor2/aqi_calculator_v1.ipynb

Описание прочих файлов, расположенных в данной директории:

- 1) AirBase_v8_stations.csv - метаданные AirBase
- 2) final_metadata.csv - итоговые метаданные, полученные объединением метаданных PanEuropean_metadata.csv и AirBase_v8_stations.csv
- 4) kml_elements.csv - вручную размеченные шаблоны для экспорта информации о станциях в kml-формате
- 5) 2022_all_stations.kml - актуализированная информация об измерительных станциях для Google Earth, включающая в себя координаты, код станции и тип станции.
- 6) denmark_data.zip - датасет по измерениям загрязнителей Дании после этапа EDA
- 7) ts_final_dk.csv - timeseries 1 станции Дании полученный в aqi_calculator_v1.ipynb, и используемый в файле ml_experiments_dk.ipynb