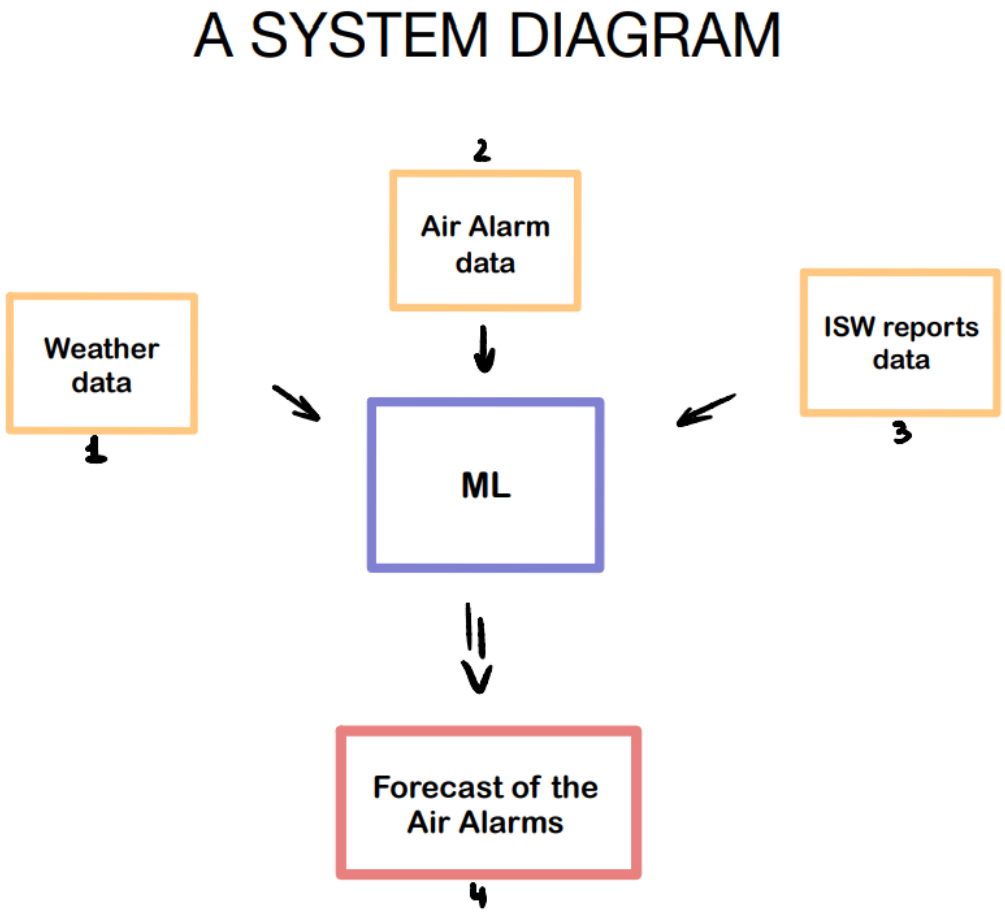
Report. Homework 2



NOTE: all the data gathered from the period 2022-02-24 - 2023-01-25

1. Weather data gathered from the given period
2. Data on the air alarms from the past period
3. Data from the ISW website
4. Output of the module with prediction of the air alarm on hourly basis

ML - machine learning model which from the given preprocessed data executes the forecasting model

Part 2:

1. To get all the data for the time period 2022-02-24 - 2023-01-25
2. Preprocess all ISW reports data

- create a pandas DataFrame to store the data
- get rid of unnecessary information such as authors' names, links etc.
- prepare functions
- tokenize the data
- filter stop words
- stemming and lemmatizing
- use TF-IDF to split the text into vector of words and calculate the inverse data frequency

Team 6:

Koval Svyatoslav

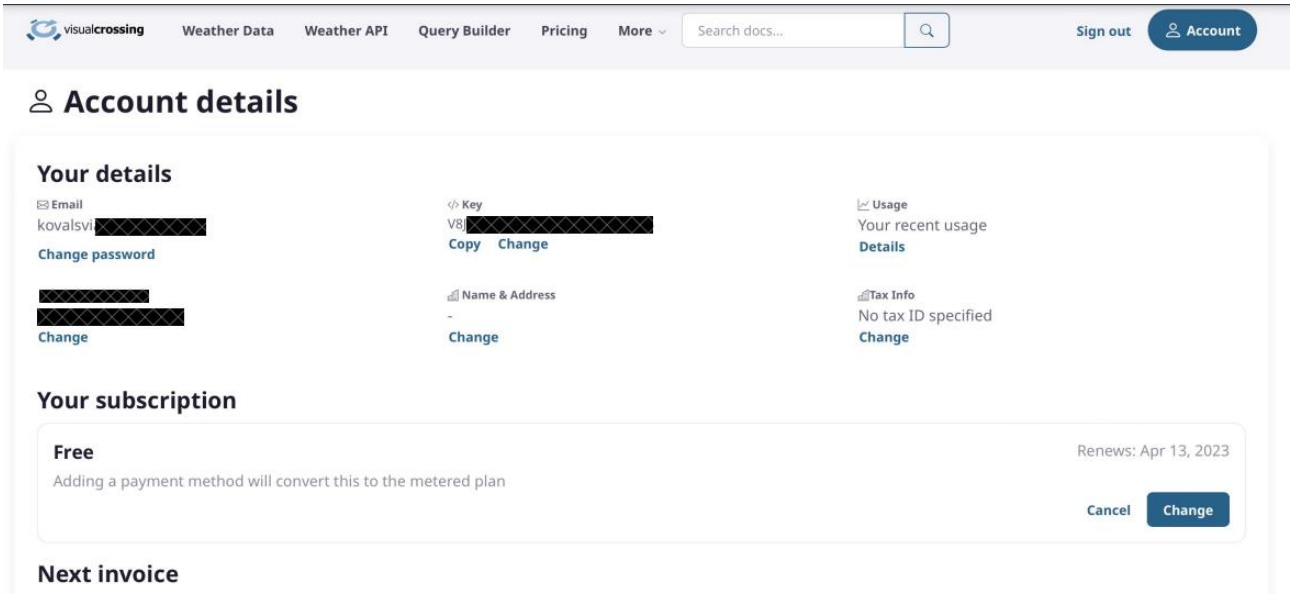
Kotliarenko Anastasiia

Semenets Daryna

Link on github: https://github.com/kovalsviat/naukma_coursework_team6

Screens

Weather API



Weather

```
df_weather.head()
```

	city_latitude	city_longitude	city_resolvedAddress	city_address	city_timezone	city_tzoffset	day_datetime	day_datetimeEpoch	day_tempmax	day_tempm
0	49.4407	32.0637	Черкаси, Україна	Cherkasy,Ukraine	Europe/Kiev	2.0	2022-02-24	1645653600	4.9	-0
1	49.4407	32.0637	Черкаси, Україна	Cherkasy,Ukraine	Europe/Kiev	2.0	2022-02-24	1645653600	4.9	-0
2	49.4407	32.0637	Черкаси, Україна	Cherkasy,Ukraine	Europe/Kiev	2.0	2022-02-24	1645653600	4.9	-0
3	49.4407	32.0637	Черкаси, Україна	Cherkasy,Ukraine	Europe/Kiev	2.0	2022-02-24	1645653600	4.9	-0
4	49.4407	32.0637	Черкаси, Україна	Cherkasy,Ukraine	Europe/Kiev	2.0	2022-02-24	1645653600	4.9	-0

5 rows × 67 columns

Alarms

```
df_alarms.head()
```

	id	region_id	region_title	region_city	all_region	start	end	clean_end	intersection_alarm_id
0	2	3	Вінниччина	Вінниця	0	2022-02-25 22:55:42	2022-02-25 23:41:53	2022-02-25 23:41:53	NaN
1	4	12	Львівщина	Львів	0	2022-02-26 06:26:17	2022-02-26 07:15:28	2022-02-26 07:15:28	NaN
2	5	14	Одещина	Одеса	0	2022-02-26 07:16:58	2022-02-26 07:47:03	2022-02-26 07:47:03	NaN
3	6	6	Житомирщина	Житомир	0	2022-02-26 08:05:54	2022-02-26 09:36:36	2022-02-26 09:36:36	NaN
4	7	3	Вінниччина	Вінниця	0	2022-02-26 08:39:39	2022-02-26 10:42:41	2022-02-26 10:42:41	NaN

```
df_alarms.shape
```

```
(19933, 9)
```

Regions

```
df_regions.head()
```

	region	center_city_ua	center_city_en	region_alt	region_id
0	АР Крим	Сімферополь	Simferopol	Крим	1
1	Вінницька	Вінниця	Vinnytsia	Вінниччина	2
2	Волинська	Луцьк	Lutsk	Волинь	3
3	Дніпропетровська	Дніпро	Dnipro	Дніпропетровщина	4
4	Донецька	Донецьк	Donetsk	Донеччина	5

```
df_regions.shape
```

```
(25, 5)
```

ISW 2022-02-24 - 2023-01-25