ARDC/ UTS Project - Team 3 (Pyro Vision)

https://universe.roboflow.com/custom-thxhn/fire-wrpgm/dataset/8

💡 Use this page to track updates on your project work (this is shared with the partner and the teaching team). For each week, add information in the sub-pages on the work done by your team and action items for follow-up. Note that the workbook is also assessed at the end (Assignment 1A).

ILAB2_IDEA.docx

Team Information

The team consists of the below members:

Student n**ame (coordinator in bold)**	Email	Background and skills	Why this project?
Nutan Thapa	**Nutan.Thapa@student.uts.edu.au**		**The project is an open-source project and if the problem is chosen wisely then it has the capability of having an impact on people's lives.**
Shivani Nandkishor Nipane	**ShivaniNandkishor.nipane@student.uts.edu.au**		**The project is focusing on crisis situations means that the work could have a tangible positive impact on communities. This alignment with social good can be personally fulfilling and resonate to make a difference.**
Ashutosh Jayant Patil	**AshutoshJayant.Patil@student.uts.edu.au**		**This open source initiative aims to address significant problems which possess the potential to improve lives. It's primary focus on crisis situations is in line with the social good and provides an opportunity to be personally fulfilled and make a difference. Collaboration improves crisis management and creates a knowledgeable network devoted to life-saving solutions.**
Ahmed Khursheed	**Ahmed.Khursheed@student.uts.edu.au**	**Advanced Mathematical Knowledge Problem Solving abilities Programming &**	**The collaborative effort in this crisis related concerned project not only advances crisis management but also**

Student n**ame (coordinator in bold)**	Email	Background and skills	Why this project?
		Computational skills Critical thinking Collaboration.	**creates a skilled network dedicated to impactful, life-saving solutions.**

Team Name: ARDC- Team 3

Sub-pages

👣 Weekly updates

Week 12 (23 October 2023 - 29 October 2023)

The Final week.

The final report is over, and every member is writing their individual exploration part. The report will be submitted by October 24. Interim Presentation

The public GitHub contains all notebook with codes and dataset use in project

GitHub Repo link : https://github.com/ahmedkhursheed/PyroVision_Greece_Wildfire

The final report file (without individual exploration and Appendix heading)

Report_ARDC_Wildfire_team 3.pdf

The group is working on making PowerPoint presentation slides for the final presentation and making changes to posters. By October 26, the PowerPoint presentation slides and PowerPoint will be updated on this page for the assessment of A1 Part A: Group Progress Tracker for project goals.

The final PPT and Poster ppt is completed. The group is ready for giving their final presentation on October 30.

These are final ppt and poster file.

Final Presentation.pptx

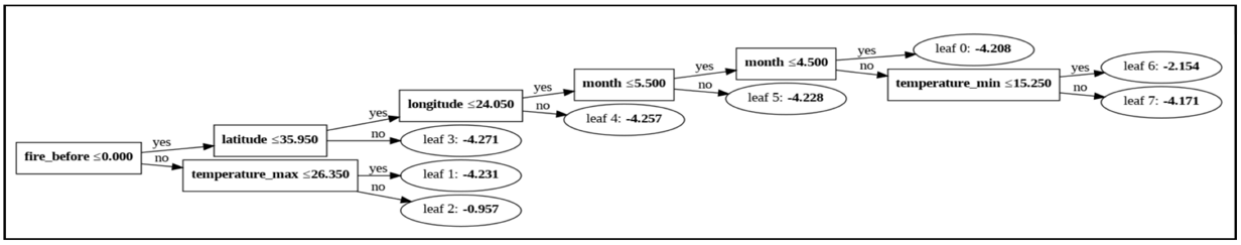
Team_3(PyroVision)_Poster.pptx

That's all update for A1 Part A: Group Progress Tracker for project goals.

Week 11 (16 October 2023 - 22 October 2023

Prediction Model

A decision tree from the trained LightGBM model was visualized, providing insights into how the model makes predictions.

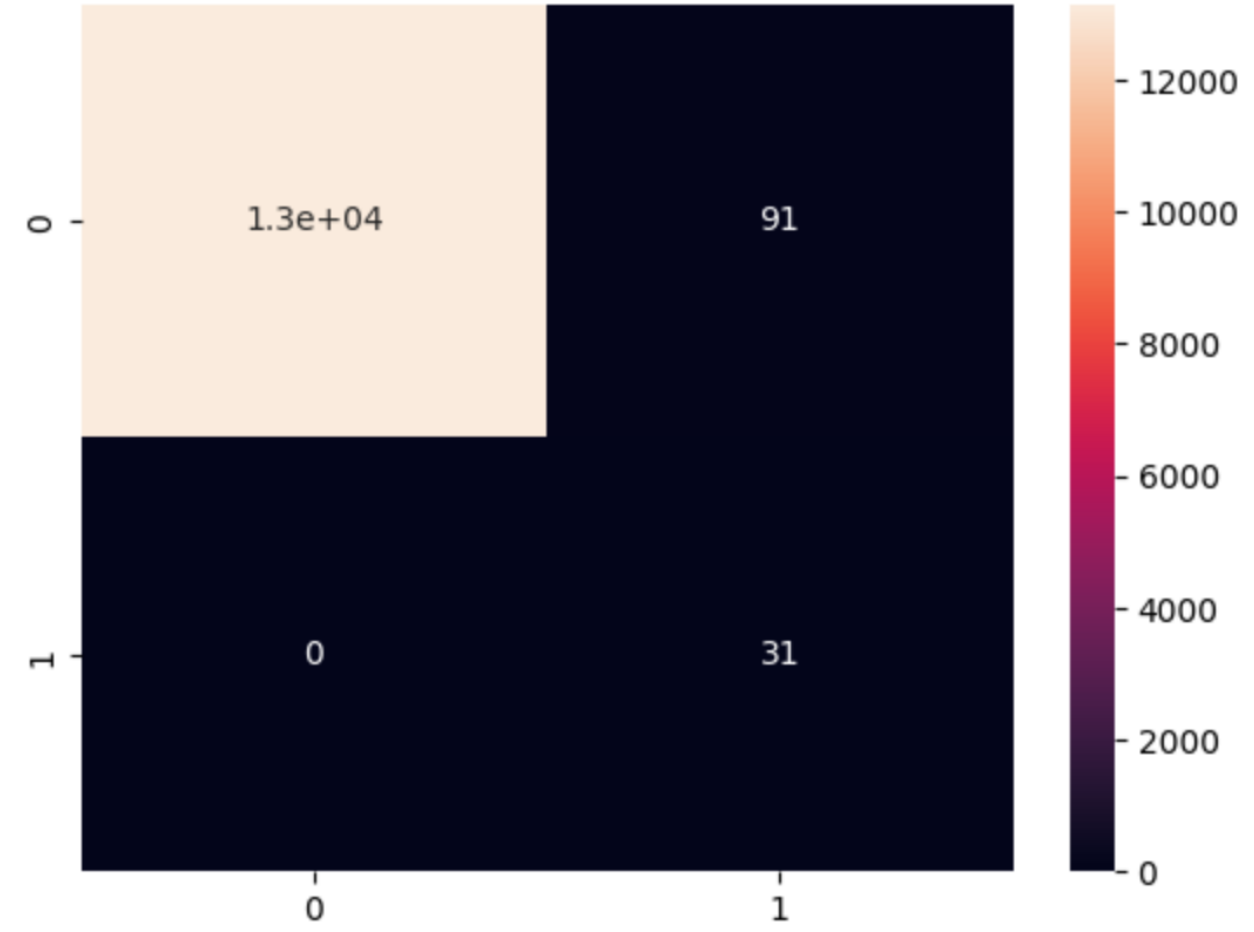


Model implementation

	precision	recall	f1-score	support
not fire	1.00	0.99	1.00	13240
fire	0.25	1.00	0.41	31
accuracy			0.99	13271
macro avg	0.63	1.00	0.70	13271
weighted avg	1.00	0.99	1.00	13271

```
[ ]  test_auc = roc_auc_score(test.fire, test_predictions)
     test_auc
```

```
0.9965634441087613
```



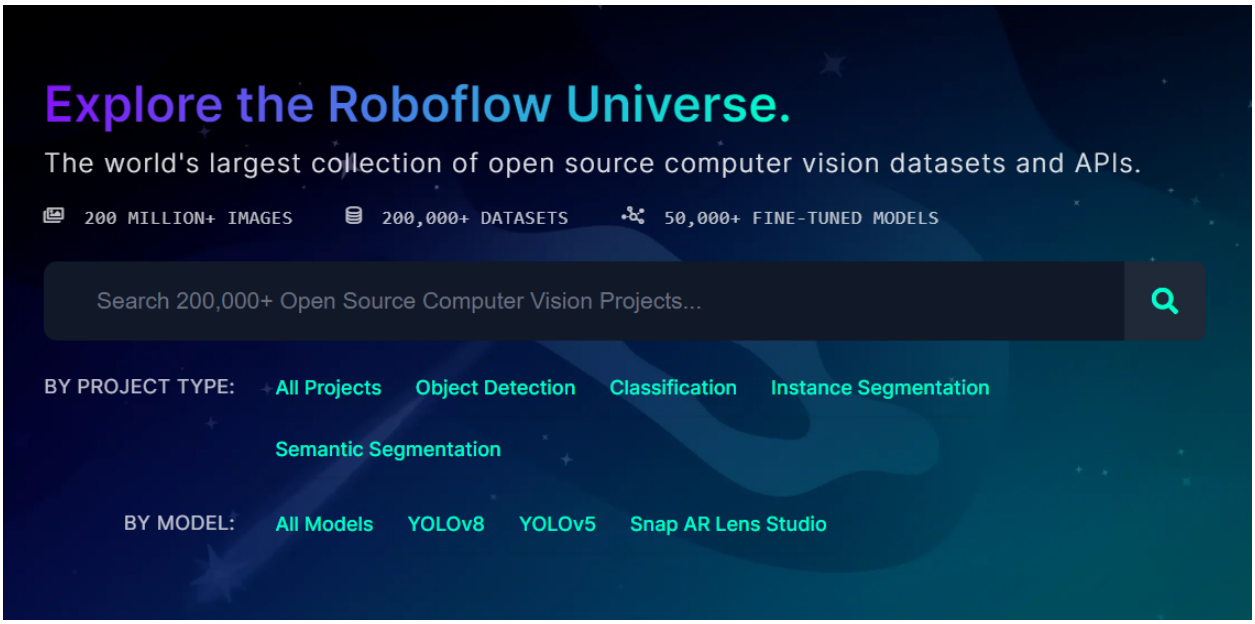
Final Report is completed.

But the optional part which is stream lit is still in progress if it will be done by 23 October then we will add it in report and submit the report.

Week 10 (09 October 2023 - 15 October 2023)

We initially attempted to create a smoke and fire detection system using the Ultralytics YOLOv5 model. However, we encountered issues related to library versions that hindered our progress. As a result, we decided to transition to YOLOv8, an improved version of the object detection model.

In order to use the YOLO model effectively, we needed to prepare our dataset, including the annotation of data. To streamline this process, we leveraged Roboflow, a platform that offers annotation tools and allows us to easily download the annotated dataset. Roboflow also provides access to a wide range of open-source datasets, enhancing the diversity of data available for our project.



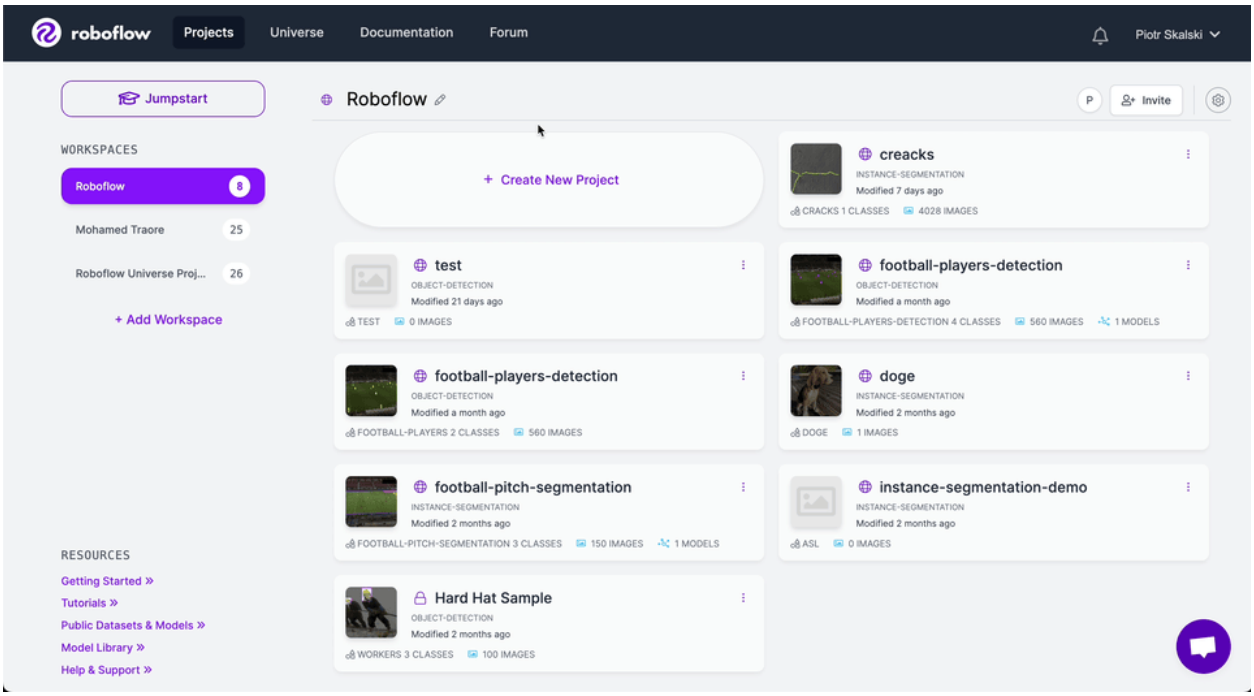
https://universe.roboflow.com/

Preparing a Custom Dataset

Building a custom dataset can be a time-consuming process, often taking dozens or even hundreds of hours to collect, label, and export images in the proper format. Fortunately, Roboflow simplifies this process, making it efficient and straightforward. Let me guide you through it!

Step 1: Creating a Project

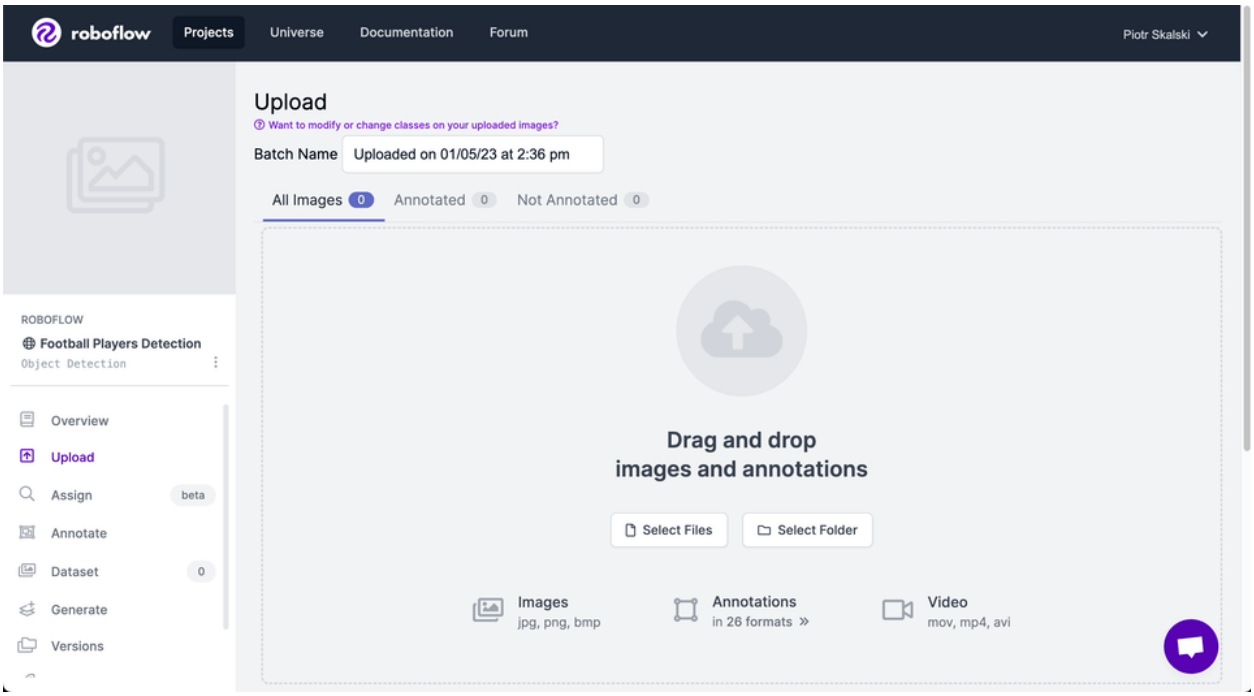
Before you begin, you'll need to create a Roboflow account. Once you've done that, you can create a new project in the Roboflow dashboard. Be sure to select the correct project type; in our case, it's Object Detection.



Step 2: Uploading Images

Next, add your data to the newly created project. You can do this via API or by using our web interface.

If you drag and drop a directory containing a dataset in a supported format, the Roboflow dashboard will automatically read the images and annotations together.

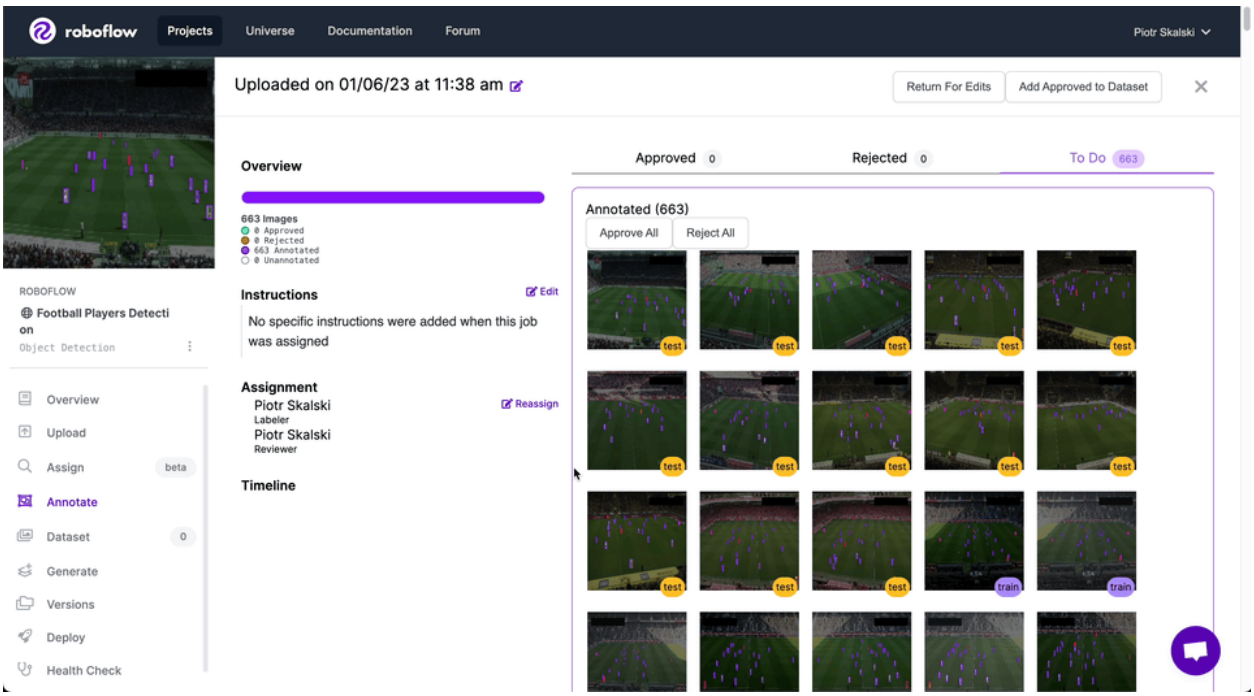


Step 3: Labeling

If you only have images, you can label them using Roboflow Annotate.

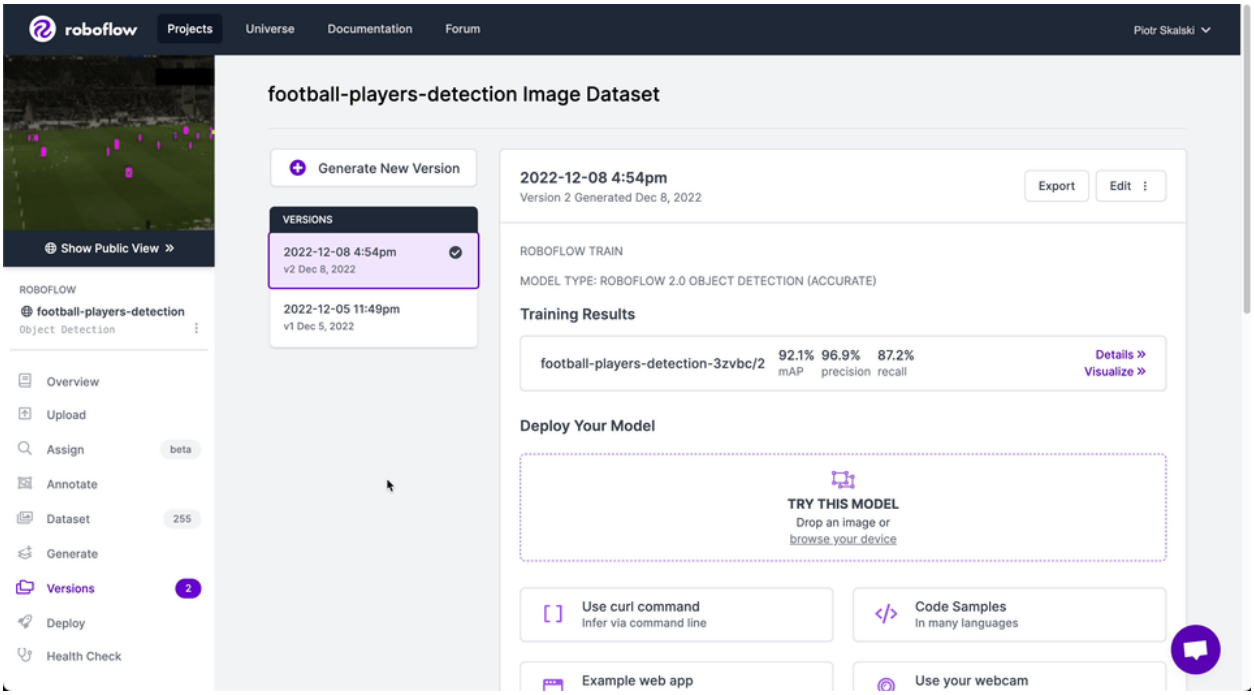
Step 4: Generate a New Dataset Version

Now that we have our images and annotations added, we can generate a Dataset Version. When generating a Version, you have the option to add preprocessing and augmentations. This step is entirely optional but can significantly improve your model's robustness.



Step 5: Exporting the Dataset

Once the dataset version is generated, we have a hosted dataset that we can load directly into our notebook for easy training. Click on `Export` and select the `YOLO v5 PyTorch` dataset format.



After the dataset is ready we can train our model on both images and videos. Our model performance is decent and is able to work on both images and videos.

Unstructured_Datasets(Image)

> https://prod-files-secure.s3.us-west-2.amazonaws.com/3db05873-554e-4332-8f59-33f279bb9c70/5eb05959-da1d-4f47-aabc-c070f90cdee6/predict.mp4

Week 9 (02 October 2023 - 08 October 2023)

- We final got temperature datasets from different stations and coordinates from multiple sites.
- The first site we requested data from was:

https://berkeleyearth.org/data/

But the Greece temperature we received from here was not good enough. Maybe others countries temperature request can be more better.

- The second site to scrape the dataset is European Climate Assessment and Dataset.

https://www.ecad.eu/

On this site, after reading the information, there is a data section, which will open the page:

Blended ECA dataset

Daily maximum temperature TX	Sources	Stations
Daily minimum temperature TN	Sources	Stations
Daily mean temperature TG	Sources	Stations
Daily precipitation amount RR	Sources	Stations
Daily mean sea level pressure PP	Sources	Stations
Daily cloud cover CC	Sources	Stations
Daily humidity HU	Sources	Stations
Daily snow depth SD	Sources	Stations
Daily sunshine duration SS	Sources	Stations
Global radiation QQ	Sources	Stations
Daily mean wind speed FG	Sources	Stations
Daily maximum wind gust FX	Sources	Stations
Daily wind direction DD	Sources	Stations

Non-blended ECA dataset

Daily maximum temperature TX	Sources	Stations
Daily minimum temperature TN	Sources	Stations
Daily mean temperature TG	Sources	Stations
Daily precipitation amount RR	Sources	Stations
Daily mean sea level pressure PP	Sources	Stations
Daily cloud cover CC	Sources	Stations
Daily humidity HU	Sources	Stations
Daily snow depth SD	Sources	Stations
Daily sunshine duration SS	Sources	Stations
Global radiation QQ	Sources	Stations
Daily mean wind speed FG	Sources	Stations
Daily maximum wind gust FX	Sources	Stations
Daily wind direction DD	Sources	Stations

The files will be pretty big search data as per the project requirement, and we will delete the unnecessary information.

- The third site which we use for the Greece case is:

https://weatherandclimate.com/greece#citytemper

This site has Greece's state and city temperatures on a monthly basis.

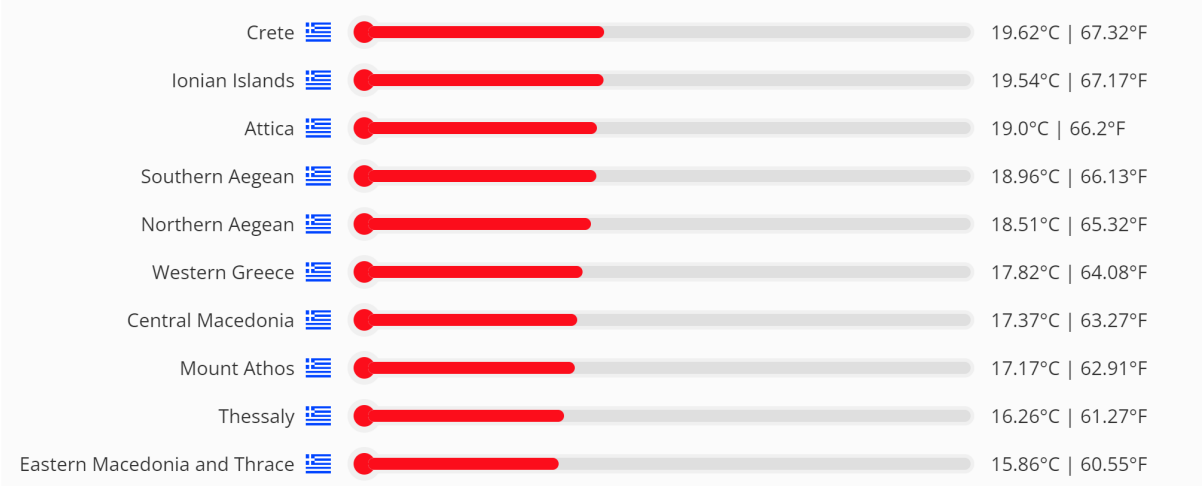
Climate Greece: Weather By Month

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Nov	Oct	Dec	Year
Record high °C (°F)	14.09 (57.36)	14.86 (58.75)	16.11 (61.0)	19.23 (66.61)	23.54 (74.37)	28.27 (82.89)	31.95 (89.51)	33.06 (91.51)	27.66 (81.79)	23.0 (73.4)	19.21 (66.58)	15.84 (60.51)	33.06 (91.51)
Average high °C (°F)	9.59 (49.26)	11.33 (52.39)	13.89 (57.0)	17.45 (63.41)	21.75 (71.15)	26.05 (78.89)	29.03 (84.25)	29.6 (85.28)	25.61 (78.1)	20.33 (68.59)	15.93 (60.67)	11.48 (52.66)	19.34 (66.81)
Daily mean °C (°F)	7.44 (45.39)	9.21 (48.58)	11.64 (52.95)	15.07 (59.13)	19.37 (66.87)	23.72 (74.7)	26.58 (79.84)	26.97 (80.55)	23.17 (73.71)	18.1 (64.58)	13.74 (56.73)	9.26 (48.67)	17.02 (62.64)
Average low °C (°F)	4.5 (40.1)	5.88 (42.58)	7.45 (45.41)	9.88 (49.78)	13.68 (56.62)	17.66 (63.79)	20.27 (68.49)	20.9 (69.62)	18.12 (64.62)	14.22 (57.6)	10.7 (51.26)	6.45 (43.61)	12.48 (54.46)
Record low °C (°F)	-1.64 (29.05)	0.44 (32.79)	2.39 (36.3)	5.1 (41.18)	8.32 (46.98)	12.22 (54.0)	15.19 (59.34)	16.15 (61.07)	12.96 (55.33)	8.94 (48.09)	5.03 (41.05)	0.36 (32.65)	-1.64 (29.05)
Average precipitation mm (inches)	65.16 (2.57)	50.76 (2.0)	49.05 (1.93)	38.21 (1.5)	38.54 (1.52)	35.65 (1.4)	21.88 (0.86)	16.31 (0.64)	31.62 (1.24)	39.67 (1.56)	50.53 (1.99)	62.35 (2.45)	41.64 (1.64)
Average precipitation days (≥ 1.0 mm)	8.73	8.59	8.86	6.78	8.03	7.78	4.21	3.64	5.39	5.97	6.55	7.98	6.88
Average relative humidity (%)	75.89	75.98	72.91	68.97	66.91	63.54	58.56	56.32	62.08	68.37	74.21	74.5	68.19
Mean monthly sunshine hours	7.18	8.56	10.19	12.25	13.73	14.17	14.18	13.62	11.47	8.89	7.81	7.37	10.78

- Even for cities, there is data available that can be used:
-

Greece's Cities by Temperature

Greece's states/regions ranking by average yearly temperature:



- Then still some data was missing and invalid, which was taken from different sites. Every site doesn't have all cities or stations One of the other sites used is

https://weatherandclimate.com/greece/central-greece

Central Greece Climate Summary

Located at an elevation of 158.53 meters (520.11 feet) above sea level, Central Greece has a Mediterranean, hot summer climate (Classification: Csa). The city's yearly temperature is 15.44°C (59.79°F) and it is -1.58% lower than Greece's averages. Central Greece typically receives about 39.61 millimeters (1.56 inches) of precipitation and has 74.01 rainy days (20.28% of the time) annually.

Country	Greece
Longitude	22.7152131
Latitude	38.6043984
Attitude/Elevation	158.53m (520.11ft)
Local time	Saturday 06:42
Annual high temperature	17.94°C (64.29°F)
Annual low temperature	10.5°C (50.9°F)
Average annual precip.	39.61mm (1.56in)
Warmest month	August (28.61°C / 83.5°F)
Coldest Month	January (2.04°C / 35.67°F)

Here data was available for central Greece:

Historical Data

Summary 2020 April Go

Temperature	Max	Average	Min	
Max Temperature	20.0°C (68.0°F)	13.7°C (56.66°F)	4.0°C (39.2°F)	
Avg Temperature	18.0°C (64.4°F)	11.47°C (52.65°F)	4.0°C (39.2°F)	
Min Temperature	13.0°C (55.4°F)	5.87°C (42.57°F)	-2.0°C (28.4°F)	
Dew Point	Max	Average	Min	
Dew Point	8.0°C (46.4°F)	4.6°C (40.28°F)	0.0°C (0°F)	
Precipitation	Max	Average	Min	Sum
Precipitation	70.9mm | 2.79in	6.09mm | 0.24in	0.0mm | 0in	182.7mm | 7.19in
Snowdepth	0.0mm | 0in	0.0mm | 0in	0.0mm | 0in	0.0mm | 0in
Wind	Max	Average	Min	

The csv file is ready.

Greece_temperature.csv

The wildfire dataset is already shared and discussed: But sharing again;

Wild_Fire_Data.csv

The prediction model is ready still trying to improve some performance level to reach AUROC at least above 0.96. By next week which is October 15, the GitHub repo which contain all notebook related to model will be shared.

Week 8 (25 September 2023 - 01 October 2023)

After researching we got the temperature dataset from this website (https://www.ecad.eu/)for a Greece one substation. We tried to look into the dataset and performed some EDA on the temperature dataset and got some visualization based on wildfire count on minimum temperature, maximum temperature and average temperature based on the different month and year and also did some visualisation based on this substation latitude and longitude. But the still there was some issue with the merging of wildfire and temperature dataset as due to latitude and longitude of the both datasets which we need to figure out. We also did cleaning and pre-processing of the wildfire dataset and got some visualisation on wildfire count based on latitude and longitude, and also tried to visualize the wildfire count based on the month and year.

Temperature Merged Dataset.xlsx

Cleaned Dataset

Wild_Fire_Data.csv

Temp_Data.csv

Code

Wild_Fire_Data.ipynb

temperature dataset.ipynb

Data Analysis.ipynb

Week 7 (18 September 2023 - 24 September 2023)

Mid-point presentations and AT3A Interim Presentation submission.

The presentation went well with proper feedback on 18. There was valuable feedback to improve our work. Then slides were submitted on Canvas as part of interim presentation submissions.

The only goal now was to find a temperature dataset because, after receiving that, our work of searching for data will be over and the only thing left is to execute our project plan.

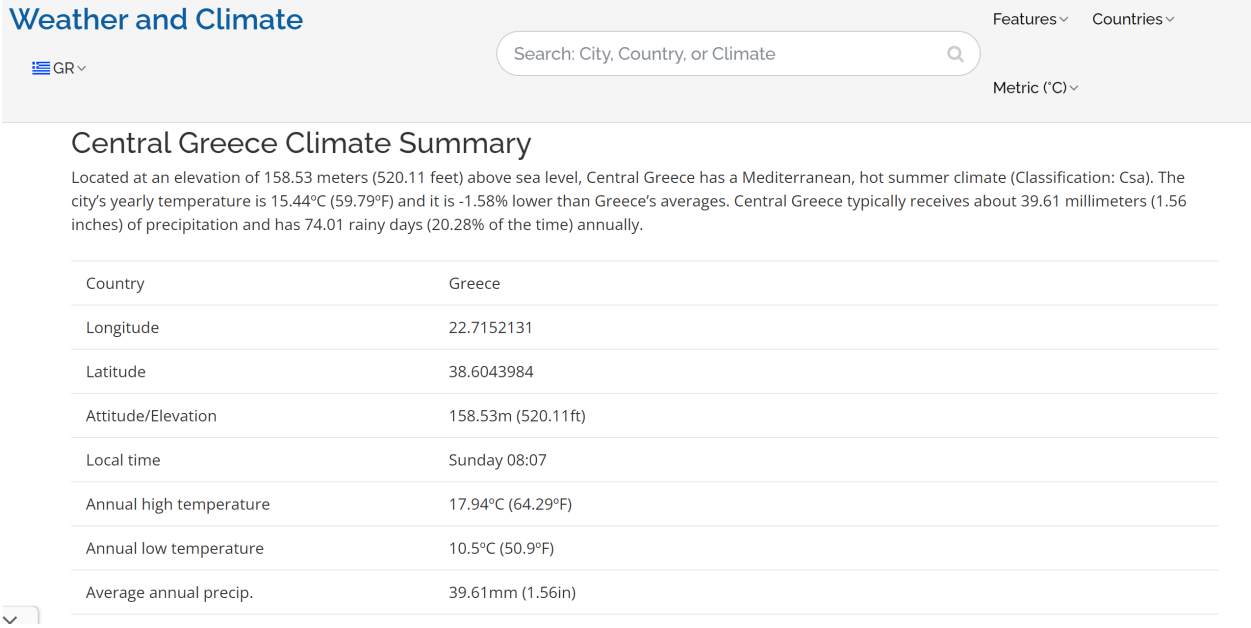
The temperature data we receive from https://berkeleyearth.org/data/ was finally received and we made an csv file of this data from year 2013 - 20223.

	A	B	C	D	E	F	G	H	I	J
1	Year	Month	Monthly Anomal	Uncertainity	Annual Anomaly	Uncertainity	5 Year Anomaly	Uncertainity	10 Year Anomaly	Uncertainity
2	2013	1	1.762	NaN	1.765	NaN	1.261	NaN	1.23	NaN
3	2013	2	1.654	NaN	1.752	NaN	1.265	NaN	1.233	NaN
4	2013	3	1.824	NaN	1.691	NaN	1.27	NaN	1.236	NaN
5	2013	4	2.108	NaN	1.481	NaN	1.274	NaN	1.239	NaN
6	2013	5	2.846	NaN	1.454	NaN	1.277	NaN	1.242	NaN
7	2013	6	1.034	NaN	1.462	NaN	1.281	NaN	1.245	NaN

The problem with this dataset is that there are too many Nan values, and the formula to calculate is applicable for the years 1980–1998. After those years, no proper information is provided. So currently, we are in a difficult situation to deal with this.

We found one more site to find temperature dataset which is

https://tcktcktck.org/greece/central-greece



This site has temperature dataset we want with addition to wind speed as well which is the second best external factor after temperature. But the problem is that the wildfire data is based on latitudes and longitudes and this is based on cities. So we need to make changes in one of them. But we are still working on this and we hope it will be solve soon.

Week 6 (11 September 2023 - 17 September 2023)

Project Work

1. **Mid-Point Peer Review in SPARK:**
   This week, our primary task was to complete the mid-point peer review in SPARK. The objective was to provide honest feedback on the contributions of our team members. Unlike group work, this review allowed each team member to express their individual views and assessments. There was no formal discussion within the group about this, but the emphasis was on honesty and constructive feedback.

2. **Focus on Temperature Dataset:**
   Another significant focus for the week was awaiting the temperature dataset. This dataset is crucial for our project, and we were eagerly anticipating its arrival. Our project's progress relies heavily on this data, and it was a key element of our work during this period.
3. **Preparation of Presentation Slides:**
   Alongside the anticipation for the temperature dataset, we were actively engaged in preparing slides for our upcoming presentation. The name of the team was Pyro Vision because pyro means fire and vision means forecasting which was our first goal of project. The initial task was crafting the cover page of these presentation slides, which marked the starting point of our week's work.

4. **Preparation of Presentation Slides**

There were instructions on what details should be available in the slides, which were

- project background
- team roles
- data and technique used
- project-finding outcomes
- achievements so far
- and the next steps (future work).

All these have to be included, and the presentation should be 15-20 minutes. Make this 5 point Write this properly.

5. **Completion of presentation slide**

By the end of the week, we successfully completed a comprehensive set of presentation slides comprising 19 pages. Recognizing the importance of a well-structured presentation, we also diligently crafted a script to accompany the slides. This script was deemed essential to ensure a smooth and effective delivery of our presentation within the given time frame.

Team 3 ARDC Wild fire Presentation (AT3A).pptx

Mid Point Presentation Script..pdf

Week 5 (04 September 2023 - 10 September 2023)

Project Phase 3: Data Analysis and Modelling

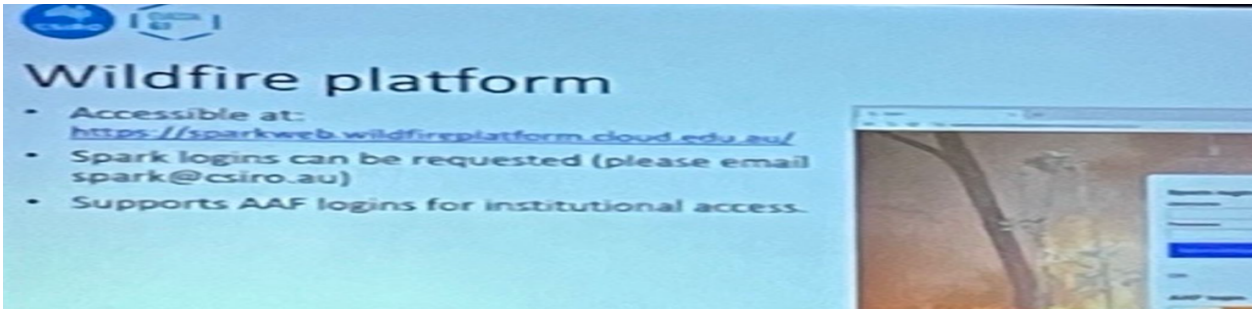
The email regrading the issue we all were facing was already sent to the tutors, and we got a reply to discuss this at the next team meeting. In the mean time, we all decide to make a backup plan with some other crisis topic rather than ozone, which was already giving us so many troubles, floods, and storms, because these 2 topics were already discussed by tutors with teams 1 and 2. So we all need a new topic now.

During week 4, we investigated the impact of Australian bushfires on the depletion of the ozone layer. Consequently, we have all previously read numerous articles and papers on this subject. And for the majority of the meeting, our instructors took a keen interest in and provided guidance regarding this particular bushfire-related crisis. Since there is still time to alter the topic and give it a fresh start, we all agreed that we could attempt to spread the topic like wildfire. This time, however, we will not fix the topic and create a new scoping document until we identify an approachable concept and the necessary dataset.

After the meeting, there were links shared by the tutors that we thought might be of any help:

- https://crisisnlp.qcri.org/
- https://datasource.kapsarc.org/explore/dataset/saudi-arabia-coronavirus-disease-covid-19-situation/information/?disjunctive.daily_cumulative&disjunctive.indicator&disjunctive.event&disjunctive.city_en&disjunctive.region_en

Now the hunt for datasets and some old work in this field that can give us any idea has started. The same day after meeting Doctor Gnana sent a message that held the promise of valuable data for anyone working on bushfires. The message mentioned that we could request data from "Spark".



The photo might have been slightly blurred, but the link was unmistakable, as provided below:
https://sparkweb.widfireplatform.cloud.edu.au/

After this we do further investigation and finally found a site where we can request for dataset related bushfire year wise. The link of that site is given below:

https://firms.modaps.eosdis.nasa.gov/download/

The datasets of different years were requested from this site, and within 20 minutes, most of the dataset files were received.

Id	Source	Area of Interest	Request Date	Status	Delete
380378	SUOMI VIIRS C2	2022-01-01 : 2023-01-01 Greece	2023-09-09 06:30:25	**Processed** on 2023-09-09 07:52:56 **Download**	
380369	SUOMI VIIRS C2	2015-01-01 : 2016-01-01 Greece	2023-09-09 04:44:55	**Processed** on 2023-09-09 07:06:57 **Download**	
380368	J1 VIIRS C1	2015-01-01 : 2016-01-01 Greece	2023-09-09 04:44:55	**Processed** on 2023-09-09 05:56:06 Download - N/A	
380367	MODIS C6.1	2015-01-01 : 2016-01-01 Greece	2023-09-09 04:44:55	**Processed** on 2023-09-09 05:56:05 **Download**	
380363	SUOMI VIIRS C2	2014-01-01 : 2015-01-01 Greece	2023-09-09 04:42:47	**Processed** on 2023-09-09 06:32:24 **Download**	

So now we have a proper dataset available for bushfires, and we need a temperature dataset in order to develop a model that can predict bushfire occurrence. But the country in which we all worked had yet to decide what we would be doing for the unstructured part.

After a few researches, we finally came to the decision to go with "Greece" because of the new image shared on the group.


Greece wildfire declared largest ever recorded in EU

Eleven planes and helicopter from bloc sent to tackle fire that has burned more than 300 sq miles of land

We finally found one link where we could also ask for a temperature dataset, which was

https://berkeleyearth.org/data/

We requested the Greece temperature dataset, which is still in progress.

So by the end of this week, we will finally have a plan with 90% of the dataset available to us.

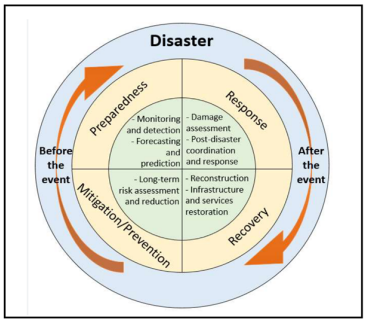
So now in this project, we will be doing the following tasks:

- Creating a prediction model for predicting wildfire occurrences in Greece
- Smoke detection in a photo or video of a recent wildfire incident in Greece
- Visualising area burned by wildfire on different years in Greece along with the estimated cost to suppress wildfires per fire in Greece
- At last, if we have enough time left, we will try to make a stream data app relevant to our project.

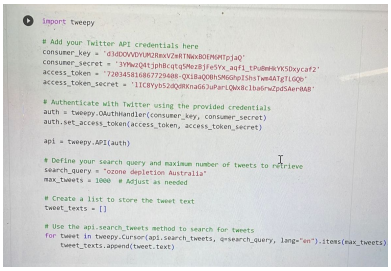
So, we will submit our new scoping document, and we hope that we will get the recent temperate dataset that we want for our project as quickly as possible.

Week 4 (28 August 2023 - 03 September 2023)

One month was about to be over, and the topic and dataset were still unclear to every group member. After reading so many articles and Google Scholar-published papers, all we could find was that ozone is healing and that further working on this specific crisis will not be as beneficial to us as we are expecting. Because the initial plan of the team was based on this principle:



Now if predictions show no depletion in ozone, then this plan of action for working on this crisis will come to nothing. Another problem we were all facing was unstructured datasets. At the start of this subject, our tutors set us the challenge of getting both structured and unstructured datasets of our preference. We were told that we could use our machine learning algorithm to extract Twitter and meta data. So our hope for a dataset was primarily based on this dataset, which we were going to obtain from social media platforms.

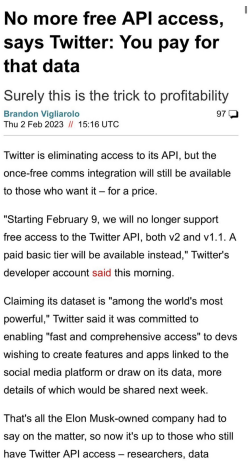


```
import tweepy

# Add your Twitter API credentials here
consumer_key = 'd3dDOVVVDYUMURmxVZmRTNWxBOEMGMTpjaQ'
consumer_secret = '3YMwzQ4tjphBcqtqSMezBjFe5YX_aqfi_tPuBmHkYK5Dxycaf2'
access_token = '720345810867729408-QXI8aQO0h5MGGhpIShsTwm4ATgTLQOo'
access_token_secret = 'IICBYyh52dQdRKna66JuPaMlQNx8clba6rwZpdSAmr0AB'

# Authenticate with Twitter using the provided credentials
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth)

# Define your search query and maximum number of tweets to retrieve
search_query = "ozone depletion Australia"
max_tweets = 1000  # Adjust as needed

# Create a list to store the tweet text
tweet_texts = []

# Use the api.search_tweets method to search for tweets
for tweet in tweepy.Cursor(api.search_tweets, q=search_query, lang="en").items(max_tweets):
    tweet_texts.append(tweet.text)
```

So after watching a few videos on YouTube, everyone of us tried to learn how we could scrape Twitter data.

But unfortunately, no one out of four members got even one percent success in this mission. Then we started finding out why these codes didn't work and found one news story about which we were all completely unaware.

Then we read an article which explained to us why we are failing.



No more free API access, says Twitter: You pay for that data

Surely this is the trick to profitability

Brandon Vigliarolo 97
Thu 2 Feb 2023 // 15:16 UTC

Twitter is eliminating access to its API, but the once-free comms integration will still be available to those who want it – for a price.

"Starting February 9, we will no longer support free access to the Twitter API, both v2 and v1.1. A paid basic tier will be available instead," Twitter's developer account said this morning.

Claiming its dataset is "among the world's most powerful," Twitter said it was committed to enabling "fast and comprehensive access" to devs wishing to create features and apps linked to the social media platform or draw on its data, more details of which would be shared next week.

That's all the Elon Musk-owned company had to say on the matter, so now it's up to those who still have Twitter API access – researchers, data

This article drained all our hopes; everything was planned on the basis of the dataset that we will obtain from Twitter. So we all finally made a decision to ask all the respective tutors how to approach this topic further because, as of now, everyone of us feels that this particular topic for this project is a dead end if we don't have any proper dataset. So we finally mailed them and told them all the problems we were facing.

Week 3 (21 August 2023 - 27 August 2023)

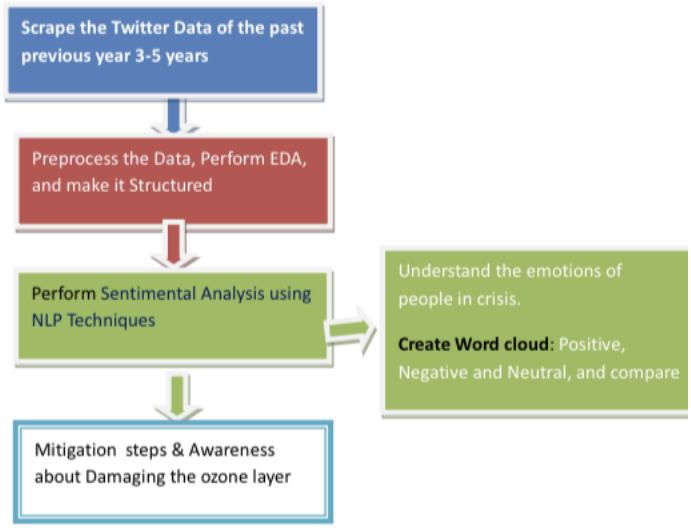
Earlier, we were thinking about or trying to work on **investigating the impact of bushfires on ozone layer depletion in Australia**, as we were trying to find some correlation between ozone layer detection and bushfires. We explored data on bushfires and the ozone layer. However, we could not find many relevant datasets indicating a direct correlation between ozone depletion and bushfires. Bushfire is not the major cause of ozone layer depletion; it's just a part of affecting the ozone layer. Just the bushfire in Australia has shown depletion in the ozone layer. Bushfires in the U.S. have not released any effects on the depletion of the ozone layer due to bushfires.

We feel that by adding the bushfire to ozone layer depletion, we are restricting our research and not exploring other factors that are actually contributing to ozone layer depletion. We think that ozone layer depletion is a serious crisis that needs to be looked at and mitigated, as in the long run it would cause some serious health issues.

We want to work on **ozone layer depletion,** as many people are not aware of and concerned about ozone layer depletion and how it could become a serious crisis in the future.

The two types of data we want to work on are unstructured and structured data.

1) Unstructured data: We want to use Twitter data to analyse how people reacted to ozone layer depletion caused by bushfires in 2019. The figure shows the basic idea that we want to implement.



Firstly, we will scrape the Twitter data for the last 3–5 years.

Secondly, we will pre-process the data to clean it.

Thirdly, we are planning to do some **sentiment analysis using NLP techniques** to see people's emotions about this crisis situation. We will try to make three word clouds based on people's responses—positive, neutral, and negative—and compare their sentiments as there is much awareness about this crisis or not and how they feel and react to it. We will also try to give mitigation

steps on how to stop damaging the environment and spread awareness among the people about how they are directly or indirectly affecting the environment, which could cause so many health issues.

2) Structured data: We are planning to use the previous year's data to first analyse the ozone data to see how the health was before and after which year the ozone layer started to deplete. We want to do **time series analysis** on the ozone layer data to see the pattern and try to **predict the health of the ozone layer.** We also want to check which substances are actually contributing to ozone layer depletion using **machine learning algorithms.** This would give clarity on what substances we need to stop using that are directly or indirectly depleting the ozone layer. This would help us understand what measures and mitigations we need to take in the long term.

ILAB_STRUCTURED DATA.docx

Week 2 (14 August 2023 - 20 August 2023)

After the first on campus discussion with our partner, we all decided to see whether we could find some correlation between bushfire and ozone layer depletion. We named it **Investigating the impacts of bushfires on ozone layer depletion in Australia** . We started to collect more evidence investigating about this problem. Luckily we found some data on bushfire and ozone layer depletion data but due to limited data we just explored more data that could give us a better correlation. So, we created the draft what can we can implement if we found the relevant datasets.

ILAB2_IDEA.docx

ILAB2_EXECUTE.docx

Useful links on bushfire.docx

Week 1 (7 August 2023 - 13 August 2023)

After researching, our teammates have decided to go with the topic, which is “Ozone layer depletion in Australia”. Australia’s ozone layer has been depleted over years and a bushfire in 2020 has also impacted the ozone layer. So we want to analyse the reasons for the ozone layer in Australia and its impact on the people of Australia leading to skin cancer issues.

We got some datasets from the Australian government website to analyse the ozone layer situation in past years

The links which we have gathered are as follows:

1)https://www.dcceew.gov.au/environment/protection/ozone/publications/antarctic-ozone-hole-summary-reports

2)https://stories.ecmwf.int/monitoring-harmful-uv-rays-and-the-ozone-layer/index.html

3)

australian-global-emissions-ozone-depleting-substances.pdf

18197.pdf

Sounds good. The key is to make sure the project you do conforms to Open Science.

Also, some alignment with Crisis is needed. As long as you can focus on this aspects, that is great.

Key resources

Project Details

Partner engagement dates

Name	Date	Tags
Check-in meeting (Date TBD)	@August 22, 2023 → August 25, 2023	
Partner meet and greet	@August 14, 2023	
Mid-point presentation	@September 18, 2023	
Final presentation	@October 30, 2023	
Check-in meeting (Date TBD)	@September 5, 2023 → September 8, 2023	
Check-in meeting (Date TBD)	@October 10, 2023 → October 13, 2023	

Project planning

Projects

Project name	Status	Owner	Dates	Priority	Completion	Blocked By
Getting started with the project	In Progress		@August 7, 2023 → August 11, 2023	High		
📈 Sample Task: Defining Project directions	Planning		@August 14, 2023 → August 18, 2023	Medium		Getting started with the project
🍏 Defining outcomes, set milestones with partner	Planning		@August 14, 2023 → August 14, 2023	Medium		📈 Sample Task: Defining Project directions
New Project	Backlog		@April 11, 2024 → April 24, 2024			