EML Research

According to the EDPB, EML (Emotionally Manipulative Language) or emotional steering, uses words or visual elements (such as style, colours, pictures or others) in a way that conveys information to users in either a highly positive outlook, making users feel good, safe or rewarded, or a highly negative one, making users feel anxious, guilty or punished. The way the information is presented to users influences their emotional state in a way that is likely to lead them to act against their data protection interests.

Sentiment analysis is the process of analysing digital text to determine of the emotional tone of the message is positive, negative, or neutral. This is according to AWS (Amazon Web Services).

Originally, sentiment analysis was going to be used to detect EML in tweets. It was advised to build a large dataset of tweets first for training and testing on a machine learning model. An amount of €100 was paid for access to the X/Twitter API by a group member allowing access to 10,000 tweets which were pulled by code to an Excel sheet. Aside from this, there was a 7-year-old Kaggle dataset of 1.6 million tweets and an NRC Word-Emotion Association Lexicon of 14,000 words and 8 emotions. Aside from the API, it would have been important to email the creators of both data sources to let them know the specific way their data was going to be used. Having addressed the API and two data sources, there was another way datasets of tweets were obtained, which was from authors of research papers.

**Figure 1. Code to pull, display & download tweets from X/Twitter API (In comment)**

import requests

import pandas as pd

import os  # Import the os module

from google.colab import files  # Import for downloading files

# Replace this with your actual Bearer Token

BEARER\_TOKEN = 'AAAAAAAAAAAAAAAAAAAAAAV8wAEAAAAApmJ8ItjlvYaUA3yvECSI9l048L8%3D9bSTWxNGdJzBrb8u9BhzTPEdzDPLyubvc9z0fUhvLduRKMVz0A'

# Function to create headers for authorization

def create\_headers(bearer\_token):

    headers = {

        "Authorization": f"Bearer {bearer\_token}",

        "Content-Type": "application/json"

    }

    return headers

# Function to search tweets by hashtag

def search\_tweets\_by\_hashtag(hashtag, max\_results=10):

    search\_url = "https://api.twitter.com/2/tweets/search/recent"

    # URL-encode the hashtag (use %23 to represent the # symbol)

    query = f"%23{hashtag} lang:en"

    # Define the parameters for the search

    query\_params = {

        'query': "elections",  # Hashtag search

        'max\_results': 100,  # Maximum number of tweets to retrieve (between 10 and 100)

        'tweet.fields': 'created\_at,author\_id,text,public\_metrics',  # Optional: Specify additional fields

    }

    # Send the GET request

    headers = create\_headers(BEARER\_TOKEN)

    response = requests.get(search\_url, headers=headers, params=query\_params)

    # Check for a successful request

    if response.status\_code == 200:

        return response.json()  # Return the tweet data in JSON format

    else:

        print(f"Error: {response.status\_code}, {response.text}")

        return None

#=====

# Function to save tweets to an Excel file

def save\_tweets\_to\_excel(tweets, filename='tweets.xlsx'):

    if tweets and 'data' in tweets:

        # Create a list of dictionaries for each tweet

        tweet\_data = []

        for tweet in tweets['data']:

            tweet\_data.append({

                'Tweet ID': tweet['id'],

                'Author ID': tweet['author\_id'],

                'Created At': tweet['created\_at'],

                'Tweet': tweet['text'],

                'Retweets': tweet['public\_metrics']['retweet\_count'],

                'Likes': tweet['public\_metrics']['like\_count']

            })

#====

# Create a DataFrame from the list of dictionaries

        df = pd.DataFrame(tweet\_data)

# Save the DataFrame to an Excel file

        df.to\_excel(filename, index=False)

        print(f"Tweets saved to {filename}")

        # Download the file to your local device

        files.download(filename)  # This should be after saving the file

    else:

        print("No tweets found to save.")

#=====

# Example usage: Search for tweets with the hashtag #USElections

tweets = search\_tweets\_by\_hashtag("USElections", max\_results=5)

#======

# Print and save the fetched tweets

if tweets:

    save\_tweets\_to\_excel(tweets)

#=====

# Print the fetched tweets

if tweets and 'data' in tweets:

    for tweet in tweets['data']:

        print(f"Tweet ID: {tweet['id']}")

        print(f"Author ID: {tweet['author\_id']}")

        print(f"Created at: {tweet['created\_at']}")

        print(f"Tweet: {tweet['text']}")

        print(f"Retweets: {tweet['public\_metrics']['retweet\_count']}, Likes: {tweet['public\_metrics']['like\_count']}")

        print("-" \* 50)

Eleven authors were contacted to get permission to use their datasets and to get their professional opinion on the project as well. Four out of the eleven responded and gave permission to use their datasets and had no problem giving advice on the project. A problem was present where all eleven emails were sent through a school Outlook account. Due to the domain name of some of the emails, Outlook refused to send them.

**Figure 2. X/Twitter API Limit**

A white rectangular object with a black line

Description automatically generated with medium confidence

I found that aside from code, sentiment analysis could be done in an Excel sheet. This required the use of the Microsoft Azure Add-in in Excel, but it couldn’t work. It was either internet error messages that would appear or because of permissions set on school emails, it was made impossible to download add-ins.

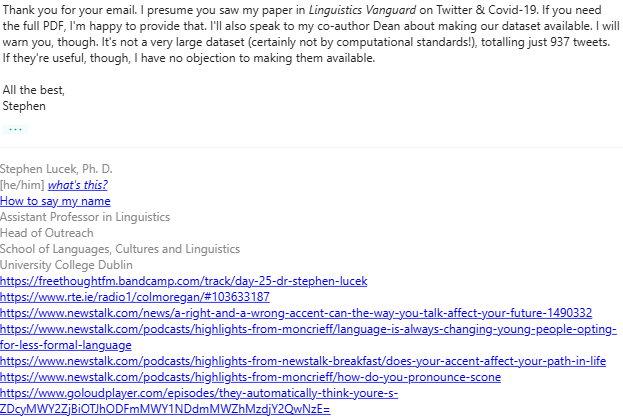
**Figure 3. 4 out of eleven authors who emailed**

A screenshot of a computer

Description automatically generated

A white background with black text

Description automatically generated



A close-up of a letter

Description automatically generated

**A screenshot of a computer

Description automatically generatedFigure 4. Sentiment Analysis** **Excel** **Errors**A screenshot of a computer error

Description automatically generated

It was eventually realized that construction of the large dataset of tweet’s would have to cease. It had proven to be too time consuming to make a large dataset of tweet’s that would cover the long list of EML techniques. Another reason for this decision was the use of sentiment analysis, it wasn’t possible to expand it beyond the three classifications it uses to address the long list of EML techniques.

**Articles of 11 people emailed**

**Zayed, O., McCrae, J. P., & Buitelaar, P. (2019). Crowd-Sourcing A High-Quality Dataset for Metaphor Identification in Tweets. *DROPS-IDN/v2/Document/10.4230/OASIcs.LDK.2019.10*. 2nd Conference on Language, Data and Knowledge (LDK 2019).** [**https://doi.org/10.4230/OASIcs.LDK.2019.10**](https://doi.org/10.4230/OASIcs.LDK.2019.10)

**Antypas, D., Preece, A., & Camacho-Collados, J. (2023). Negativity spreads faster: A large-scale multilingual twitter analysis on the role of sentiment in political communication. *Online Social Networks and Media*, *33*, 100242.** [**https://doi.org/10.1016/j.osnem.2023.100242**](https://doi.org/10.1016/j.osnem.2023.100242)

***Reading Twitter as a marketplace of ideas: How attitudes to COVID-19 are affecting attitudes to migrants in Ireland*. (n.d.). Retrieved October 6, 2024, from** [**https://www.degruyter.com/document/doi/10.1515/lingvan-2021-0158/html**](https://www.degruyter.com/document/doi/10.1515/lingvan-2021-0158/html)

***Sentiment Analysis on Twitter using Neural Network: Indonesian Presidential Election 2019 Dataset—IOPscience*. (n.d.). Retrieved October 6, 2024, from** [**https://iopscience.iop.org/article/10.1088/1757-899X/1077/1/012001**](https://iopscience.iop.org/article/10.1088/1757-899X/1077/1/012001)

***Twitter deleted Russian troll tweets. So we published more than 200,000 of them.* (2018, February 14). NBC News.** [**https://www.nbcnews.com/tech/social-media/now-available-more-200-000-deleted-russian-troll-tweets-n844731**](https://www.nbcnews.com/tech/social-media/now-available-more-200-000-deleted-russian-troll-tweets-n844731)