

a4-Multilingual-Large-Language-Models-dan-jang

December 1, 2023

1 CS410: Natural Language Processing, Fall 2023

1.1 A4: Multilingual Large Language Models (LLMs), Dan Jang - 11/27/2023

1.1.1 Content Warning: *Alike many “natural language processing” oriented or sources for text & data, the text & data used in this project may have been collected from public websites or sources, and may be unfiltered. Thus, some text & data may be disturbing, disagreeable, or otherwise objectionable.*

1.2 Description of Assignment

1.2.1 Introduction

As the training & testing datasets from our first & second assignments has become defunct recently (*The Multilingual Amazon Reviews Corpus* by Phillip Keung, Yichao Lu, György Szarvas, & Noah A. Smith (October 6th, 2020) [1]), in this assignment, *A4: Multilingual Large Language Models (LLMs)*, we will be using a new dataset, the “*Unified Multilingual Sentiment Analysis Benchmark*” [2] - where we will be exploring & comparing the performance of two, specific **Large Language Models (LLMs)**:

1. *Meta’s LLaMA 2* (*Large Language Model Meta AI 2*) [3], [4]

&

2. *OpenAI’s ChatGPT* (As of November 27th, 2023, **ChatGPT** is currently running the November 6th, 2023 Update model version of *GPT-3.5 Turbo gpt-3.5-turbo-1106*) [5]-[7]

Like the previous two assignment, this assignment focuses on using specific NLP models - specifically, *Meta’s LLaMA 2* [3], [4] and *OpenAI’s ChatGPT Large Language Models (LLMs)* [5], [6].

As a different approach in comparison to our previous three (3) assignments, in this current assignment, *A4*, instead of implementing a full text-classification model for sentiment-prediction, we will instead be exploring the techniques of *prompt engineering* [8] in creating a text-classifier - in comparison to previous assignment, *A3*, where we had used the monolingual *Pretrained Language Models (PLMs)* of *BERT* [9] and *GPT-2* [10].

1.2.2 Data Preparation

In *A4*, we will be using a new dataset, *Unified Multilingual Sentiment Analysis Benchmark* [2], which is a multilingual dataset with Tweets labeled with a **sentiment** value, where the classification labels are as follows:

0 indicating *negative*,

1 indicating *neutral*,

...and a 2 indicating *positive*.

The *Unified Multilingual Sentiment Analysis Benchmark* [2] dataset contains sentiment-labeled Tweets in Arabic, English, French, German, Hindi, Italian, Portuguese, and Spanish.

Instead of using several thousands of data-entries, for both a training and a testing dataset to train and test our text-classification model like we've explored in our previous assignments, of assignments *A1* through *A3* - in this assignment, *A4*, we will *instead* be utilizing a **much smaller subset of data** from our *new dataset* [2] & only utilizing a testing dataset based on a sampling of Tweet-entries across the different languages, in lieu of using both a testing and training dataset.

Specifically, we will be creating a multilingual, Tweet-entry-based testing dataset, of a recommended amount of ~50 instances (the actual number used in my implementation will be specified below in my *Text Classification Through Prompt Engineering* section) of the testing dataset for each language as aforementioned, and sentimentality-wise, we will only look at the Tweet-entries with either the *positive* or *negative* sentiment labels, and will be ignoring Tweet-entries with the *neutral* sentiment label.

Although no training will be required, we will still need to carefully prepare a balanced test set across both classes, and throughout our various languages sampled from the *new dataset* [2].

1.2.3 Prompt Engineering

Researcher Golam Md. Muktedir of the University of California, Santa Cruz, Computer Science and Engineering (muktadir@ucsc.edu) describes in good detail, the past-to-present history of *prompt engineering* in their relatively recent September 30th, 2023 paper, "*A Brief History of Prompt: Leveraging Language Models*" [8].

Modern day *prompt engineering* can be described by the following description...

"...researchers and practitioners [who] have explored various techniques to harness the full potential of language models, leveraging the power of prompts to guide, control, and tailor the output of these sophisticated AI systems..."

...where this *very description* of *prompt engineering* originates from a generated text borne of *prompt engineering* itself, as given by Researcher Muktedir in their paper's introduction, from these two prompts [8]:

"Prompt #1: *You are a scholar in machine learning and language models. I am writing a paper on the history of prompt engineering and generation. Can you give me a timeline for prompt engineering evolution? (We used this timeline to create prompts for each section later)"*

&

"Prompt #2: *Write the introduction of this paper. Emphasize that this paper focuses on how language prompts and queries have been used so far."*

While the techniques of *prompt engineering*, seemingly only recently popularized, or rather, been made the spotlight mainstream subject of discussion of recent, in regard to the wide field of *Natural Language Processing (NLP)* - this popularization, of course, significantly due to the rise of *OpenAI*'s

[ChatGPT](#) [6], [7] - however, nonetheless are a set of techniques a history long prior to this very novel advent of widely accessible platforms providing services based on *Natural Language Processing (NLP)* study & more specifically, through *Large Language Models (LLMs)* [8].

As section-titled in their paper, Researcher MuktaDir describes an early era of *prompt engineering* in “*Prehistoric Prompting: Pre NN-Era*”, which describes among the most preliminary studies into what we know as *prompt engineering* today - as far back as the 1960s through the 1970s with “*Early Natural Language Interfaces*”, the 1990s through the 2010s with “*Advances in Natural Language Processing*”, e.g. *Neural Networks (NNs)* and *Machine Learning (ML)*, to modern-day, featuring the rise of **Multilingual Large Language Models (MLLMs)** of those, we will exploring of two such **MLLMs** [8].

1.2.4 Multilingual Large Language Models (Multilingual LLMs, or ‘MLLMs’)

One way we can/will be accessing our two **Multilingual Large Language Models (MLLMs)**, [Meta’s LLaMA 2](#) [3], [4] and [OpenAI’s ChatGPT](#) **Multilingual Large Language Models (MLLMs)** [5], [6], would be through the graphical interface chat platform provided through [HuggingFace Chat](#) [11].

1.2.5 Meta’s Open-Source LLaMA 2 (*Large Language Model Meta AI 2*) Multilingual Large Language Model (MLLM) [3], [4]

The first **MLLM** we will be exploring will be the *open-source*, [LLaMA 2](#) (*Large Language Model Meta AI 2*) [4] **MLLM**, which was first co-released by [Meta](#) and [Microsoft](#), back on June 17th, 2023 [14].

This **MLLM** was first released & described as **LLaMA (Version 1)** in the paper, “[LLaMA: Open and Efficient Foundation Language Models](#)” [13] - which was first published earlier this year in February 27th, 2023, by the [Meta AI](#) team [3] - authored by Researchers Hugo Touvron, and thirteen (13) other authors.

This first version of **LLaMA** was described as “*a collection of foundation language models ranging from 7B to 65B parameters*”, these models having been trained on “*trillions of tokens*”, and importantly, claiming that *it is and was possible* to “*train state-of-the-art models using publicly available datasets exclusively*” - claiming, that the-then **LLaMA-13B** model outperformed the **GPT-3 (175B)** model on “*most benchmarks*” [13].

The second iteration, **LLaMA 2** has been released as part of various *fine-tuned LLMs*, e.g. **LLaMA 2-Chat**, **Code LLaMA 2**, etc. - as to represent the latest efforts by [Meta](#) and [Microsoft](#) to provide their updated advancements of their open-source **LLM** - where, in their [paper](#), they claim that the latest **LLaMA 2 LLM(s)** are capable and *very* comparable in performance to those of other closed-source **LLMs**, either in regard to a closed-source based set of techniques used to create the **LLM(s)** and/or a closed-source service that is provided as *SaaS (Software as a Service)*, e.g. [OpenAI’s ChatGPT](#) [14], [4], [5], [6].

1.2.6 OpenAI’s Proprietary ChatGPT (*Chat Generative Pretrained Transformer*) Multilingual Large Language Model (MLLM) [5]-[7]

The most famous **MLLM** model as of recent mainstream attention, is the proprietary model [ChatGPT](#) (based on *Generative Pretrained Transformer 3.5 Turbo*, currently on the November 6th, 2023 update-iteration, [gpt-3.5-turbo-1106](#)) is based off advancements made by [OpenAI](#) on the

closed-source ‘release’ of **GPT-3**, and of course, follows the progress made from the second predecessor, **GPT-2** [6], [7], [5], [15], [10].

Recalling from *A3*: In **GPT-2**’s [technical paper](#), *OpenAI* researchers Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever describes the **GPT-2** model - a **uni-directional** model trained on a dataset of *8 million web pages* with **1.5 billion parameters** (*OpenAI*, Radford et al. [10]).

However, if we were to make a quick GPT lineage comparison between **gpt-3.5-turbo-1106** vs. **GPT-2**, of course, the former, **gpt-3.5-turbo-1106**, has up to **175 billion parameters** & the latter, at least the largest-parameter version of **GPT-2**, has still only **1.5 billion parameters** - representing an extremely significant increase in parameter-power in **gpt-3.5-turbo-1106** over its long predecessor **GPT-2** model.

1.2.7 Comparison of LLaMA-2 vs. ChatGPT Multilingual Large Language Models (MLLMs) Architectures & Designs

When comparing **LLaMA 2** vs. **ChatGPT (gpt-3.5-turbo-1106)**, the first difference we can note is that, since **LLaMA 1** was released & now with **LLaMA 2**, **LLaMA** has remained available as a series of **open-source Multilingual Large Language Models (MLLM)**, as released & created by *Meta’s AI Team* (& *Microsoft* for version 2) - where anyone is able to request & receive access to freely download the model & its fine-tuned derivatives as per the nature open-source vs. the proprietary nature of **ChatGPT**, which is only offered indirectly, as a **Software-as-a-Service (SaaS)** through the free version of *OpenAI*’s main *ChatGPT web-platform* (chat.openai.com) or as a paid service through *OpenAI*’s *API platform* [4], [6], [7], [3], [13], [14], [6], [5].

From September 1st, 2023, Enterprise AI Strategist Sunil Ramlochan of the *Prompt Engineering Institute* described in the summary of their article, [How Does Llama-2 Compare to GPT-4/3.5 and Other AI Language Models](#) that **LLaMA 2** “competes on accuracy despite lower complexity [vs. the complexity of GPT-3.5]” [16] - where such strong accuracy could be through *Meta*’s various improvements from **LLaMA 1** to **LLaMA 2**, e.g. *Ghost Attention*, which may improve “*dialogue context tracking*” [16], [4], [14].

As cited in [16], Researcher Waleed Kadous of *anyscale* performed an analysis of our two **Multilingual Large Language Models (MLLMs)**, in their supportively titled (of the claims made in [16]) article, “*Llama 2 is about as factually accurate as GPT-4 for summaries and is 30X cheaper*” [17]:

Model	Input Tokens	Output Tokens	Price MTokens input (\$)	Price MTokens output (\$)	Total cost
gpt-4	125902	834	30	60	\$3.83
gpt-3.5-turbo	125902	770	1.5	2	\$0.19
Llama-2-7b	149238	80525	0.25	0.25	\$0.06
Llama-2-13b	149238	104947	0.5	0.5	\$0.13
Llama-2-70b	149238	61500	1	1	\$0.21

1.2.8 Text Classification Through Prompt Engineering

To perform our text-classification task on these, we will **follow these steps**:

1. Prepare a balanced test set of a specific, constant chosen-size of test Tweet-instances for each of the eight different languages from our new dataset, *Unified Multilingual Sentiment Analysis Benchmark* [2]. Specifically, we need to pick either ‘positive’ or ‘negative’ classes, ergo, ignoring Tweet-instances with the ‘neutral’ label.

In my implementation, I will be using a number of **twenty-four (24)** testing-instance / Tweets, per language - for a total of **192** testing-instances / Tweets (across our **eight (8)** languages).

2. Using Prompt Engineering, we will have each of the **Multilingual LLMs** attempt to predict the sentiment, neither ‘positive’ nor ‘negative’, of each of the Tweet-instances in our test set, for each language.

1.2.9 Questions to Answer Through Results & Analysis

We will, of course, be comparing the two **Multilingual LLMs** by the metrics such as the *F1-score* or other relevant metrics like from our previous assignments we’ve explored this lovely Fall term.

Specific to this assignment, *A4*, we will be focusing on **answering the following questions** through our exploration of the two aforementioned **MLLMs** - and the results and analysis from our exploration:

1. “How do the two LLMs perform? Which one is better? Any possible explanation?”
2. “Comparing the results across the six different languages, what do you observe? Any possible explanation?”
3. “What challenges did you face?”

1.3 A4 - Libraries & Constants Initialization

```
[1]: ##### CS410: Natural Language Processing, Fall 2023 - 11/27/2023
##### A4: Multilingual Large Language Models (LLMs), Dan Jang - Initializations:
      ↪ Libraries, Models, Data, & Constants

### 0.) Libraries
from sklearn.naive_bayes import GaussianNB
# from sklearn.linear_model import LogisticRegression
# from sklearn.metrics import accuracy_score
# from sklearn.metrics import confusion_matrix
from sklearn.metrics import f1_score
# from sklearn.metrics import classification_report
# from sklearn.model_selection import train_test_split
# from transformers import AutoTokenizer, AutoModel
import torch
import json# as jn
import pandas
import os
import huggingface_hub
```

```
import numpy# as np
#import random
```

1.4 A4 - The “Multilingual And Normalization Generator from Original Set (M.A.N.G.O.S.)” Function

1.4.1 The M.A.N.G.O.S. function was borne out of a desire to remain sane whilst having properly Python-parsed testing instances for each of our eight (8) language-based datasets, heh.

```
[8]: ##### CS410: Natural Language Processing, Fall 2023 - 11/30/2023
##### A4: Multilingual Large Language Models (LLMs), Dan Jang - "Multilingual
↳And Normalization Generator from Original Set (M.A.N.G.O.S.)"
##### Just a cute name for a function that will parse .jsonl files properly, as
↳for some reason, Visual Studio Code does not have any sorting extensions
↳compatible with .jsonl files directly - at least for the raw formats of our
↳new dataset, hehe.

### A4: 1.) Multilingual And Normalization Generator from Original Set (M.A.N.G.
↳O.S.) Function

## A4: 1a.) Path & files to pre-processed testing instance-sets
datapath = 'data/tweet_sentiment_multilingual/data/processed/processedv2'
# tweetsets = [
#     'processed_arabic.jsonl',
#     'processed_english.jsonl',
#     'processed_french.jsonl',
#     'processed_german.jsonl',
#     'processed_hindi.jsonl',
#     'processed_italian.jsonl',
#     'processed_portuguese.jsonl',
#     'processed_spanish.jsonl'
# ]
langlist = [
    'arabic',
    'english',
    'french',
    'german',
    'hindi',
    'italian',
    'portuguese',
    'spanish'
]

print("Multilingual And Normalization Generator from Original Set (M.A.N.G.O.S.
↳) Function is now loading...")
```

```

# A4: 1b.) Main data array for holding the mangos (parsed .jsonl files ->
↳ Python data arrays)
mangos = {lang: [] for lang in langlist}

# A4: 1c.) The main M.A.N.G.O. parser loop
for lang in langlist:
    lepath = os.path.join(datapath, f'processed_{lang}.jsonl')

    with open(lepath, 'r', encoding='utf-8') as mango:
        # peel = line; mango = file; e.g., for each line in current
        ↳ processed_lang.jsonl file, etc.
        for peel in mango:
            tweet = jn.loads(peel)
            if tweet["label"] == "1":
                print("uh oh, we got a mango with no taste! (a neutral tweet
↳ detected, whoopsie)")
            else:
                mangos[lang].append({
                    'tweet': tweet["text"],
                    'actual-sentiment': tweet["label"],
                    'chatgpt-sentiment': None,
                    'llama2-sentiment': None,
                })

# A4: 1d.) Box all the freshly peeled mangos into a nice & tidy, new Python
↳ file for implementation
with open('tweety_mangos.py', 'w', encoding='utf-8') as candiedmangos:
    for lang, tweet in mangos.items():
        instancename = f'parsed_{lang}'
        #outfile.write(f'{list_name} = {json.dumps(data, ensure_ascii=False,
↳ indent=4)}\n\n')
        candiedmangos.write(f'{instancename} = {jn.dumps(tweet,
↳ ensure_ascii=False, indent=4)}\n\n')
        print(f'{instancename} has been successfully parsed...')

print("Multilingual And Normalization Generator from Original Set (M.A.N.G.O.S.
↳ ) Function has successfully parsed all the tweet-datasets!")

```

Multilingual And Normalization Generator from Original Set (M.A.N.G.O.S.)

Function is now loading...

parsed_arabic has been successfully parsed...

parsed_english has been successfully parsed...

parsed_french has been successfully parsed...

parsed_german has been successfully parsed...

parsed_hindi has been successfully parsed...

parsed_italian has been successfully parsed...

parsed_portuguese has been successfully parsed...

parsed_spanish has been successfully parsed...

Multilingual And Normalization Generator from Original Set (M.A.N.G.O.S.)

Function has successfully parsed all the tweet-datasets!

1.5 A4 - Prompt Engineering, Section I: ChatGPT Examples

To perform the sentiment analysis, a series of similar, *engineered prompts* like shown below were used to get the Tweet-testing instance's text data sentiment analyzed by *ChatGPT*.

Specifically, the prompt below was used to complete parts of each list of our parsed language testing-datasets, where the resultant output from prompts engineered were designed to produce a specifically formatted output, where we wanted *ChatGPT* to both perform sentiment analysis, and additionally, fill-in the sentiment rating within our parsed Python list format.

Since *ChatGPT* does seem to have a tangible - but, at least a reasonably wide - *per-request* generation-output limit on its output messages (& where you have to click **Continue Generating**) - for each parsed language data list, e.g. `parsed_arabic`, multiple prompts were used & resultant generated outputs combined cumulatively - **until, and for all**, Tweet-instances in each of our eight (8) language-based datasets were successfully analyzed by *ChatGPT* & values for sentiment prediction rating were given *ChatGPT*.

The *engineered prompt* resembled the following:

Prompt #1 - Me: "...

I need you to perform sentiment prediction as a text classification model.

Fill in the "chatgpt-sentiment" values for each item in `parsed_arabic`, based your own analysis. Specifically, fill in either a value of "0" for negative sentiment or a value of "2" for positive sentiment.

Here is the data for `parsed_arabic`:

```
```py
parsed_arabic = [
 {
 "tweet": ") (http via @user",
 "chatgpt-sentiment": None,
 "llama2-sentiment": None
 },
 {
 "tweet": " " ,
 "chatgpt-sentiment": None,
 "llama2-sentiment": None
 },
 {
 "tweet": " http",
 "chatgpt-sentiment": None,
 "llama2-sentiment": None
 },
 #... so on yes, until tweet/item[23]
 {
```



```

 "tweet": " ... http",
 "chatgpt-sentiment": None,
 "llama2-sentiment": None
 }
]
"""

```

Response - *ChatGPT*: "...

Sure, I can help you with sentiment prediction. Please note that sentiment analysis can be sub

```

parsed_arabic = [
 {
 "tweet": ") (http vía @user",
 "chatgpt-sentiment": 2,
 "llama2-sentiment": None
 },
 {
 "tweet": " " ,
 "chatgpt-sentiment": 2,
 "llama2-sentiment": None
 },
 {
 "tweet": " http",
 "chatgpt-sentiment": 2,
 "llama2-sentiment": None
 },
 {
 "tweet": "@user "\ ""\ ,
 "chatgpt-sentiment": 2,
 "llama2-sentiment": None
 },
 {
 "tweet": " - "\: ... http",
 "chatgpt-sentiment": 0,
 "llama2-sentiment": None
 },
 {
 "tweet": " ! _ # @user _ # @user _ # @user " _
 "chatgpt-sentiment": 2,
 "llama2-sentiment": None
 },
 {
 "tweet": " - - - - # "
 "chatgpt-sentiment": 0,
 "llama2-sentiment": None
 },
]

```

```
#... so on yes, until tweet/item[23], or usually, a few of the last tweets/items would be :
..."
```

Although, it might be more long-term beneficial to simply create a set of functions, such that just generates a 24-element, one-line JSON line that contains all twenty-four (24) Tweet-instances per given language out of our eight (8) language testing-sets, I used the above method for the sake being able to more readily/visually verify that *ChatGPT* (or for *LLaMA 2*) is, hopefully not, generating mismatched sentiments or nonsensical sentiments.

E.g. if some very obviously polarized Tweet or multiple super obviously one-sided Tweet-items, where the sentiment should be super on-the-nose, e.g. "tweet": "Trying to have a conversation with my dad about vegetarianism is the most pointless infuriating thing ever #caveman ", being set to "2" as somehow a positive sentiment; if such occurs, I'd be able to hopefully spot out errors.

However, besides that aspect, the aforementioned is a thought of possible long-term streamlining the process in regards to this specific task of Tweet sentiment analysis through *prompt engineering*.

## 1.6 A4 - Prompt Engineering, Section II: LLaMA 2 Examples

*LLaMA 2*, interestingly, was actually easier to work with & get more in-depth results from - yet, this clear advantage was massively overshadowed by an unfortunate aspect that the specific Chat-based instance of *LLaMA 2* to be used in this assignment kept on freezing, becoming unresponsive even with only a moderate to small input given-data in various kinds of engineered prompts, and has been completely down, service-wise, for the most of my attempts to use / check if it was back up, which heavily delayed my progress for my *A4* assignment - although, still, I apologize for my own delays nonetheless.

As a compromise, due to *HuggingChat* being unavailable for multiple hours at the time, I instead explored a specific version of *LLaMA-2*, the *LLaMA-2-7b-chat* model [19], which, was also hosted on & one of '*HuggingFace*'s *huggingface-projects* [20], but of course, the [*LLaMA-2-7b-chat*] specific version, quantitatively has *ten-fold* less parameters than the model used by *HuggingChat*, which is/was *meta-llama/llama-2-70b-chat-hf* [21] had offered & the lack of file-upload functionality is slightly a bummer.

In that thought, initially, my *engineered prompts* for *LLaMA 2* were much grander & had more input-text per prompt, e.g.:

**Prompt #2a** - Me: "...

Could you please perform sentiment analysis for each item in the following Python language test:

In each of these elements, there is a "tweet" key with a value correlating to a random Tweet message.

Specifically, provide & set the "llama2-sentiment" key-value to either a "0" for negative sentiment or a "1" for positive sentiment.

```
```py # pretend this represents an uploaded file of the parsed Python data lists, full_tweet_data,
parsed_arabic = [
    ...
]
```

```
parsed_english = [
    ...
]
```

```
#... etc.
```
```

```
...”
```

However, using the alternative, [LLaMA-2-7b-chat](#) model-chat [19], my *prompt engineering* trended back in resembling to the *piece-by-piece* sentiment analysis gathering strategy used for *ChatGPT* - however, I still used the same prompt as above, just that I was not able to do more than one whole language-testing instance-set of 24 Tweets at a time:

**Prompt #2b** - Me: ”...

Could you please perform sentiment analysis for each item in the following Python language test:

In each of these elements, there is a "tweet" key with a value correlating to a random Tweet message.

Specifically, provide & set the "llama2-sentiment" key-value to either a "0" for negative sentiment or a "1" for positive sentiment.

```
```py
parsed_arabic = [
    ...
]
```

Llama-2 7B Chat

This Space demonstrates model [Llama-2-7b-chat](#) by Meta, a Llama 2 model with 7B parameters fine-tuned for chat instructions. Feel free to play with it, or duplicate to run generations without a queue! If you want to run your own service, you can also [deploy the model on Inference Endpoints](#).

🔍 For more details about the Llama 2 family of models and how to use them with `transformers`, take a look [at our blog post](#).

🔧 Looking for an even more powerful model? Check out the [13B version](#) or the large [70B model demo](#).

Duplicate Space for private use

Chatbot

Could you please perform sentiment analysis for each item in the following Python language testing data-lists, where there are eight different language-data lists, each containing 24 elements.

In each of these elements, there is a "tweet" key with a value correlating to a random Tweet message in the language of that given language testing data-lists. Using the text data of "tweet" in each of the elements, provide & set a sentiment rating, for each element based on its "tweet" text-data, in the "llama2-sentiment" key.

Specifically, provide & set the "llama2-sentiment" key-value in each element, to either a "0" for negative sentiment or a "2" for positive sentiment, based on the text-data of "tweet" of that element:

```
parsed_arabic = [  
{  
  "tweet": "http via @user", نوال الزعبي (الشاب خالد ليس عالمي) هههههههه أتفرجى على ها الفيديو يا مبتدئة",  
  "actual-sentiment": "0",  
}
```

Type a message...

Submit

Retry

Undo

Clear

”

Response #2b - *LLaMA-2-7b-chat*: "...

Sure, I can assist you in assigning sentiment scores to each tweet in the provided list using the following Python code:

Chatbot

Sure, I can assist you in assigning sentiment scores to each tweet in the provided list using the llama2-sentiment key. Here are the results:

Index	Tweet	Sentiment Score
0	نوال الزغبى (الشاب خالد ليس عالمي) هههههههههههه	0
1	نوال الزغبى لطيفه الفنانة الوحيدده اللي كل الفيديو كليبات تبعها ماتسبب تلوث بصري ولا سمعي لو صوتها اقل من عادي	2
2	http لما قالت نوال الزغبى لابقلها هاللقب فرحوا فانزها 😂😂😂 كان لازم ياخدوها اهانة مش ثابا	2
3	@user تذكرني بأغنية نوال الزغبى [عينيك كدايين]	2
4	http ...بلا تشفير- أمل حمادي ينتقد النجمة نوال الزغبى: "نتي	0
5	نوال_الزغبى#user نجوى_كرم#user فنانة لبنانية كبيرة صوتها إسطوري؟! #ماجة_الرومي جوليا_بطرس#user	2

Type a message...

Submit

Retry

Undo

Clear

...”

However, I found another alternative for accessing [meta-llama/Llama-2-70b-chat-hf](#) [21] whilst [HuggingChat](#) [11] remained down (at 503 service error code) - which was through [llama2.ai](#) [22] which luckily - supposedly - also runs the 70B-illion parameter version of *LLaMA 2* - ...but that also broke during usage.

So, I just was stuck with using *LLaMA-2-7b-chat* model-chat [19]. Just in case, I am downloading the LLaMA 2 70B model directly in the case that even this 7B parameter model-chat goes down as well.

With the *LLaMA-2-7b-chat* model-chat [19], I engineered a pretty consistent-output prompt that seemed to be effective for the piece-by-piece approach as aforementioned:

Prompt #3a - Me: ”...

Please fill in the "llama2-sentiment" key in the following elements, based on your sentiment analysis

...

...”

Example Response #3 - *LLaMA-2-7b-chat*: ”...

Chatbot

Sure! I filled in the `llama2-sentiment` column of the provided list with a value of either 0 (negative sentiment) or 2 (positive sentiment), based on the sentiment of each tweet's text data analyzed using the ChatGPT model. Here are the results:

Index	Tweet	chatgpt-sentiment	llama2-sentiment
0	"Trying to have a conversation with my dad about vegetarianism is the most pointless infuriating thing ever #caveman "	None	0
1	"@user You are a stand up guy and a Gentleman Vice President Pence "	None	2
2	"@user @user @user Looks like Flynn isn't too pleased with me, he blocked me. You blocked by Flynn too @user "	None	0
3	"i'm not even catholic, but pope francis is my dude. like i just need him to hug me and tell me everything is okay. "	None	2

Sometimes, *LLaMA-2-7b-chat* seemed to be struggling to choose between a positive or negative sentiment, as it had given ‘neutral’ scores at times, which is not valid. Ergo, I tweaked my original **Prompt #3a** as follows:

Prompt #3b - Me: "...

Please fill in the "llama2-sentiment" key in the following elements, based on your sentiment analysis. Do not return anything else.

...
...
..."

As an additional *prompt engineering* technique, I used a ‘cleaned’ version of my language testing-instance sets, such that there are no pre-existing sentiment scores from *ChatGPT*, *LLaMA 2*, or the actual sentiment scores that may possibly influence the sentiment analysis outputs.

1.7 A4 - Main Implementation: Intermodel, Multilingual F1-Score Metrics, Performance, & Analysis Function-Implementation

```
[21]: ##### CS410: Natural Language Processing, Fall 2023 - 11/27/2023
##### A4: Multilingual Large Language Models (LLMs)
#### Objective: Exploring Natural Language Processing (NLP),
#### for a text classification task, predicting whether a piece of text is
    ↪ "positive" or "negative."
#### ...focusing on two (2) Multilingual Large Language Models (MLLMs),
#### ...Meta's LLaMA 2 [4] & OpenAI's ChatGPT (gpt-3.5-turbo-1106) [6],
#### ...and using prompt engineering to perform the classification task on
    ↪ eight (8) language-based, Tweet-testing-datasets.

# 3a.) Initialize tweet-index / language-dictionary
#rosetta_tablet = {
    # 'parsed_arabic': parsed_arabic,
```

```

# 'parsed_english': parsed_english,
# 'parsed_french': parsed_french,
# 'parsed_german': parsed_german,
# 'parsed_hindi': parsed_hindi,
# 'parsed_italian': parsed_italian,
# 'parsed_portuguese': parsed_portuguese,
# 'parsed_spanish': parsed_spanish
#}

#langlist = {
rosetta_tablet = {
    'Arabic': parsed_arabic,
    'English': parsed_english,
    'French': parsed_french,
    'German': parsed_german,
    'Hindi': parsed_hindi,
    'Italian': parsed_italian,
    'Portuguese': parsed_portuguese,
    'Spanish': parsed_spanish
}

# 3b.) Tallying up the sentiment scores!
#actualsentiments = [tweet['actual-sentiment']]

# 3c.) F1-score functions!
def sentimental_values(tweets, actualsentiments):

    actual = [int(tweet['actual-sentiment']) // 2 for tweet in tweets] #_
    ↪mitosis lol
    guesses = [int(tweet[actualsentiments]) // 2 for tweet in tweets]

    return f1_score(actual, guesses)

def tally(iPad):
    llama2_f1_scoreboard = {}
    chatgpt_f1_scoreboard = {}

    for lang, tweets in iPad.items():
        llama2_f1_scoreboard[lang] = sentimental_values(tweets, ↵
        ↪'llama2-sentiment')
        chatgpt_f1_scoreboard[lang] = sentimental_values(tweets, ↵
        ↪'chatgpt-sentiment')

    idx = 1
    for lang in iPad.keys():
        #for langy in langz:

```



```

        print(f"Language #{idx}, {lang}: The Open-Sourced Challenger, Meta's
↳LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality:
↳{llama2_f1_scoreboard[lang]:.10f}")
        print(f"Language #{idx}, {lang}: The Closed-Source Champion, OpenAI's
↳ChatGPT MLLM - F1 Scores for Tweet-Sentimentality:
↳{chatgpt_f1_scoreboard[lang]:.10f}")
        print("----\n")
        idx = idx + 1

        average_llama2_f1score = sum(llama2_f1_scoreboard.values()) /
↳len(llama2_f1_scoreboard)
        print(f"Average LLaMA 2 F1-Score: {average_llama2_f1score:.10f}")

        average_chatgpt_f1score = sum(chatgpt_f1_scoreboard.values()) /
↳len(chatgpt_f1_scoreboard)
        print(f"Average ChatGPT F1-Score: {average_chatgpt_f1score:.10f}")

tally(rosetta_tablet)

# for lang, tweets in rosetta_tablet.items():
#     tally(lang, data)

```

Language #1, Arabic: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.9230769231

Language #1, Arabic: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.8000000000

Language #2, English: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.7272727273

Language #2, English: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.9600000000

Language #3, French: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.4000000000

Language #3, French: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.6086956522

Language #4, German: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.6666666667

Language #4, German: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.8800000000

Language #5, Hindi: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.6000000000

Language #5, Hindi: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.6666666667

Language #6, Italian: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.7272727273

Language #6, Italian: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.8333333333

Language #7, Portuguese: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.8571428571

Language #7, Portuguese: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.8000000000

Language #8, Spanish: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.7000000000

Language #8, Spanish: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.8000000000

Average LLaMA 2 F1-Score: 0.7001789877

Average ChatGPT F1-Score: 0.7935869565

1.8 *A4: Multilingual Large Language Models (MLLMs)* - Results and Analysis: Meta's LLaMA-2 vs. OpenAI's ChatGPT!

1.8.1 Inter-Multilingual Performance & Analysis

Let us first compare language-by-language, the performance differences of [LLaMA 2](#) [4] vs. [ChatGPT \(gpt-3.5-turbo-1106\)](#) [6], [7]:

From our *first* language-Tweet results: **Arabic**:

Language #1, Arabic: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Sentimentality: 0.92307

Language #1, Arabic: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Sentimentality: 0.92307

...We see that [LLaMA 2](#) [4] is the winner, by **~0.123**.

These F1-scores for both models represent among the highest F1-scores for both models, which is interesting.

Perhaps, the high F1-score of 0.92307 seen of [LLaMA 2](#) [4] indicates it may have had more training with Arabic texts & data.

From our *second* language-Tweet results: **English**:

Language #2, English: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Ser
Language #2, English: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Ser

...We see that *ChatGPT (gpt-3.5-turbo-1106)* [6], [7] is the winner, by ~**0.233**.

Interestingly, the F1-score of 0.72727 seen of *LLaMA 2* which seems slightly odd, but maybe that is just a Western society-bias-belief kind of thing to expect.

From our *third* language-Tweet results: **French**:

Language #3, French: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Ser
Language #3, French: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Ser

...We see that *ChatGPT (gpt-3.5-turbo-1106)* [6], [7] is the 'winner', by ~**0.208** - but relatively, they both sucked.

The language of love, neither *LLaMA 2* [4] nor *ChatGPT (gpt-3.5-turbo-1106)* [6], [7] speak *at all fluently*. But, *ChatGPT (gpt-3.5-turbo-1106)*, in comparison, speaks at least with a F1-score of ~0.608.

From our *fourth* language-Tweet results: **German**:

Language #4, German: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Ser
Language #4, German: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Ser

...We see that *ChatGPT (gpt-3.5-turbo-1106)* [6], [7] is the winner, by ~**0.213**.

From our *fifth* language-Tweet results: **Hindi**:

Language #5, Hindi: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Ser
Language #5, Hindi: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Ser

...We see that *ChatGPT (gpt-3.5-turbo-1106)* [6], [7] is the winner, by ~**0.067**. However, it appears both LLMs are sort-of around the same range in terms of how close the F1-scores are to each other.

From our *sixth* language-Tweet results: **Italian**:

Language #6, Italian: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Ser
Language #6, Italian: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Ser

...We see that *ChatGPT (gpt-3.5-turbo-1106)* [6], [7] is the winner, by ~**0.106**.

From our *seventh* language-Tweet results: **Portuguese**:

Language #7, Portuguese: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Ser
Language #7, Portuguese: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Ser

...We see that *LLaMA 2* [4] is the winner, by ~**0.057**.

From our *eighth* language-Tweet results: **Spanish**:

Language #8, Spanish: The Open-Sourced Challenger, Meta's LLaMA 2 MLLM - F1 Scores for Tweet-Ser
Language #8, Spanish: The Closed-Source Champion, OpenAI's ChatGPT MLLM - F1 Scores for Tweet-Ser

...We see that *ChatGPT (gpt-3.5-turbo-1106)* [6], [7] is the winner, ~**0.100**.

1.8.2 Overall Average F1-Score Performance Analysis

Averaged F1-scores for each LLM, over F1-scores across all eight (8) language-Tweet testing instances yielded:

Average LLaMA 2 F1-Score: 0.7001789877

Average ChatGPT F1-Score: 0.7935869565

...Both in a relatively alright ballpark - with a **~0.0935** lead by **ChatGPT** (**gpt-3.5-turbo-1106**) [6], [7] - but for both **MLLMs**, through our exploration of these models in this *A4* assignment, we do see that these are **Multilingual Large Language Models**, indeed!

1.8.3 Multilingual Large Language Model (MLLM) Comparative Analysis & Discussion

It is very exciting to see an *open-sourced* model be *very much* in competition with the most *state-of-the-art, closed-source* models such as with **ChatGPT** (**gpt-3.5-turbo-1106**) [6], [7]!

That **LLaMA 2** [4] competes with **ChatGPT** (**gpt-3.5-turbo-1106**) [6], [7] at this kind of *multilingual* level & performance, especially, was really awesome to see!

1.8.4 Results and Analysis: Le Questions

1. “How do the two LLMs perform? Which one is better? Any possible explanation?”

As discussed in the introductory sections, one of the most prominent differences in these *Large Language Models* (**LLMs**) is that **LLaMA 2** [4] is an *open-source* model, while **ChatGPT** (**gpt-3.5-turbo-1106**) [6], [7] is a *closed-source* model.

Surely, performance by pursuit of profit, is not the only factor in the *slight* edge that **ChatGPT** (**gpt-3.5-turbo-1106**) [6], [7] has shown over its 100% free, open-source **LLaMA 2** [4] model - however, with more funding & profit, may have led to a much greater ability for *Research & Development* of *GPT* models over time, where speaking of, *OpenAI* has been long developing its *GPT* lineage.

2. “Comparing the results across the six different languages, what do you observe? Any possible explanation?”

I think two most interesting things noted were firstly, the Arabic prowess of **LLaMA 2** [4] over **ChatGPT** (**gpt-3.5-turbo-1106**) [6], [7], and secondly - rightfully so, the French language being the most difficult for both our **MLLMs** to predict sentiment, a hope that perhaps some things still remain just a little bit within the creativity of humans.

As briefly mentioned, I think the most obvious answer for **LLaMA 2** [4] having a edge over **ChatGPT** (**gpt-3.5-turbo-1106**) [6], [7] in the Arabic language may have been greater exposure to Arabic texts & training data. Or it could also have been a statistical fluke, too.

3. “What challenges did you face?”

Using a GUI for genuine *Natural Language Processing* (**NLP**) projects & analysis, was very surprisingly one of the biggest challenges. A limitation & an almost felt-delay of the simulated human-typing, letter-crawling as you wait for one of your 50+ prompts to finish being outputted by either GUI of **ChatGPT** or **LLaMA 2** was a bit of a challenge.

However, the biggest challenge, I think, was the unfortunate, massive delay caused due to *HuggingChat* being down - but *HuggingFace* is still & has been a wonderful, great (and very vital) resource in the realm of *NLP*, CS, & in general, a super cool resource - so I think it was a fluke, but one that was a bit significant & caused a lot of stress, as I tried to scramble a way to get our last assignment done - alas, we survive. Hopefully.

1.8.5 Prompt Engineering Challenges & Limitations

Even with great, simple, & pinpoint prompt engineering - I still saw that outputs still can & did vary wildly for no reason. However, this was only really present when using the *LLaMA-2-7b-chat* - if I had access to set the *temperature* or other settings through the API, I think such output-variational issues *could* have been mitigated slightly - yet, it was not a significant challenge.

A limitation in the literal sense, however, was that even with fairly simple text-classification or parsing-related prompts, a ratelimit or limit to the ‘*generatable output-in-one-go*’ were present in both LLM platforms, where using an automated API does seem, if we instead used, *A4* may not have been as much of a struggle - however, the struggle was worth it, I think. I can say I know a little about how *ChatGPT* works now, heheh.

2 References & Resources

2.1 Libraries, Dependencies, & Credits

[HuggingFace_hub](#)

[HuggingFace’s transformers](#)

[sklearn.metrics.f1_score](#)

[nbconvert](#)

Credits to GitHub Copilot & ChatGPT for code implementation assistance.

2.2 IEEE Formatted References

- [1] P. Keung, Y. Lu, G. Szarvas, and N. A. Smith, “The Multilingual Amazon Reviews Corpus.” arXiv, Oct. 06, 2020. doi: 10.48550/arXiv.2010.02573.
- [2] “cardiffnlp/tweet_sentiment_multilingual · Datasets at Hugging Face.” [Online]. Available: https://huggingface.co/datasets/cardiffnlp/tweet_sentiment_multilingual
- [3] “Introducing LLaMA: A foundational, 65-billion-parameter language model.” [Online]. Available: <https://ai.meta.com/blog/large-language-model-llama-meta-ai/>
- [4] “Llama 2: Open Foundation and Fine-Tuned Chat Models | Research - AI at Meta.” [Online]. Available: <https://ai.meta.com/research/publications/llama-2-open-foundation-and-fine-tuned-chat-models/>
- [5] “OpenAI Platform.” [Online]. Available: <https://platform.openai.com>
- [6] “Introducing ChatGPT.” [Online]. Available: <https://openai.com/blog/chatgpt>
- [7] “New models and developer products announced at DevDay.” [Online]. Available: <https://openai.com/blog/new-models-and-developer-products-announced-at-devday>

- [8] G. M. Muktadir, “A Brief History of Prompt: Leveraging Language Models. (Through Advanced Prompting).” arXiv, Nov. 28, 2023. doi: 10.48550/arXiv.2310.04438.
- [9] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.” arXiv, May 24, 2019. doi: 10.48550/arXiv.1810.04805.
- [10] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language Models are Unsupervised Multitask Learners”.
- [11] “HuggingChat.” [Online]. Available: <https://huggingface.co/chat>
- [12] “ Transformers.” [Online]. Available: <https://huggingface.co/docs/transformers/index>
- [13] H. Touvron et al., “LLaMA: Open and Efficient Foundation Language Models.” arXiv, Feb. 27, 2023. doi: 10.48550/arXiv.2302.13971.
- [14] “Meta and Microsoft Introduce the Next Generation of Llama.” [Online]. Available: <https://ai.meta.com/blog/llama-2/>
- [15] T. B. Brown et al., “Language Models are Few-Shot Learners.” arXiv, Jul. 22, 2020. doi: 10.48550/arXiv.2005.14165.
- [16] “How Does Llama-2 Compare to GPT-4/3.5 and Other AI Language Models,” Prompt Engineering Institute. [Online]. Available: <https://promptengineering.org/how-does-llama-2-compare-to-gpt-and-other-ai-language-models/>
- [17] “Llama 2 vs. GPT-4: Nearly As Accurate and 30X Cheaper,” Anyscale. [Online]. Available: <https://www.anyscale.com/blog/llama-2-is-about-as-factually-accurate-as-gpt-4-for-summaries-and-is-30x-cheaper?ref=promptengineering.org>
- [18] [A. Arshad, ”How to Use Llama 2 from Hugging Face to Solve Real-World Problems,” Medium. [Online]. Available: <https://medium.com/@anasarshad22/how-to-use-llama-2-from-hugging-face-to-solve-real-world-problems-f0216e32ef>](<https://medium.com/@anasarshad22/how-to-use-llama-2-from-hugging-face-to-solve-real-world-problems-f0216e32ef62>)
- [19] “meta-llama/Llama-2-7b-chat · Hugging Face.” [Online]. Available: <https://huggingface.co/meta-llama/Llama-2-7b-chat>
- [20] “Llama 2 7B Chat - a Hugging Face Space by huggingface-projects.” [Online]. Available: <https://huggingface.co/spaces/huggingface-projects/llama-2-7b-chat>
- [21] “meta-llama/Llama-2-70b-chat-hf · Hugging Face.” [Online]. Available: <https://huggingface.co/meta-llama/Llama-2-70b-chat-hf>
- [22] “Chat with Llama 2.” [Online]. Available: <https://www.llama2.ai/>

2.3 Special Thanks

A huge shoutout to our awesome & wonderful TA Bahareh G. for their patience in dealing with my delays & for having to read my often super long written analysis / my messy code!

Thanks & credits to my fellow classmate Ahmad A. for tips & suggestions on making a Python-specific data array for a more intuitive method to calculate metrics in the Jupyter Notebook.

The Eight (8) Raw A4 Pre-Processing Testing Language Tweet-Instance-Sets from the New Dataset [2]

Content Warning: *Alike many “natural language processing” oriented or sources for text & data, the text & data used in this project may have been collected from public websites or sources, and may be unfiltered. Thus, some text & data may be disturbing, disagreeable, or otherwise objectionable.*

[illegible]

```

{"text": "Trying to have a conversation with my dad about vegetarianism is the most pointless :"}
{"text": "@user You are a stand up guy and a Gentleman Vice President Pence ", "label": "2"}
{"text": "@user @user @user Looks like Flynn isn't too pleased with me, he blocked me. You blo"}
{"text": "i'm not even catholic, but pope francis is my dude. like i just need him to hug me and"}
{"text": "@user for al the crying you do about how middle America is left out-they have twice as"}

```



```

{"text": "Samsung to Bring Android 7.0 Nougat to Galaxy S6, S6 edge, Note 5, and Tab S2 - Softp"}
{"text": "Israel deems comatose Gaza man who needs treatment in West Bank a security threat. #"}
{"text": "I will go so far to say s1 of westworld isn't just good, it's brilliant. A story with"}
{"text": "Iraqi Forces set to storm 3 areas of #Mosul, #AlQahira, #alMasarif & #alAmn#MosulOps"}
{"text": "#NationalFastFoodDay Would love to live there. Chick-fil-A \ud83d\ude0d ", "label": "1"}
{"text": "Electoral College must reject Trump #rejecttrump #notmypresident #takingbackdemocracy"}
{"text": "The Reputation Doctor weighs in on Tony Romo #NFL @user joins @user on #TheMorningRun"}
{"text": "Stop #fracking #Cuadrilla persecuting land defenders through the legal system #Bankrupt"}
{"text": "So proud of way @user & #trumpTransitionteam are molding strong leadership group for"}
{"text": "@user Everyone's thinking far too short-termed. When all fossil fuels run out with no"}
{"text": "@user @user - #ScreamQueens so lucky to get to work with TK. ", "label": "2"}
{"text": "@user @user @user Why is it that today society is casually comfortable about being pr"}
{"text": "This is a big deal and a smart move by Microsoft: ", "label": "2"}
{"text": "Ben Carson for Housing & Urban Development?? \ud83d\ude10 I just can't \ud83d\ude12"}
{"text": "Digesting while watching #ScreamQueens s1 Thanksgiving & Black Friday eps. SO SO GOOD"}
{"text": "Israel's New Racism: Persecution of African Migrants in Holy Land #gaza #palestine #"}
{"text": "I've been listening to Leonard Cohen nonstop for the past month. Only to find out he"}
{"text": "@user She just didn't get them in areas were she needed them. Lots of voter suppress"}
{"text": "Listen to #NBAAwards Winner @user interview on @user ", "label": "2"}

```

3.) processed_french.jsonl (Pre-Processed by manually filtering out any jsonl objects with a “1” value for the “label” key, which would be the ‘neutral’ rating)

```

{"text": "Royal: le pr\u00e9sident n'aime pas les pauvres? \"c'est n'importe quoi\" http \u2026"}
{"text": "#press #seo: Nager vert avec les piscines \u00e9cologiques http #piscine\u00e9colog"}
{"text": "#Biodiversit\u00e9 #Gard : Des L\u00e9zards ocell\u00e9s d\u00e9truits sans autorisat"}
{"text": "Etat des lieux et analyse des m\u00e9canismes de soutien public aux \u00e9nergies ren"}
{"text": "Les esp\u00e8ces disparaissent 1.000 fois plus depuis que l'Homme existe | @user http"}
{"text": "Une entreprise allemande propose une solution \u00e9co-efficace de ferme urbaine http"}
{"text": "R\u00e9chauffement : l\u2019ONU pr\u00e9dit un v\u00e9ritable \u00e9chauffement climatique"}
{"text": "@user @user Ah Oui, les \u00e9oliennes, pourrait aussi les \u00e9loigner .. ouf Merc"}
{"text": "Info: \u00e9oliennes - Le r\u00e9sultat de l\u2019enquête publique est tomb\u00e9"}
{"text": "R\u00e9chauffement climatique: il est grand temps d'agir - http", "label": "2"}
{"text": "#GCO : les \u00e9cologistes strasbourgeois sont \"stup\u00e9faits\" - http", "label": "1"}
{"text": "J'aime ma bio parce que je mange bio et je travail dans la biologie et je roule en b"}
{"text": "Pol\u00e9mique autour de l\u2019abattage massif de bouquetins en Haute-Savoie | Eco"}
{"text": "Echange autour du th\u00e8me de l'agriculture raisonn\u00e9e \u00e0 la R\u00e9union l"}
{"text": "Un r\u00e9seau de journalistes pour lutter contre les n\u00e9gationnistes du changem"}
{"text": "@user \"Illimit\u00e9es\", cela fera sourire les environnementalistes. Et le th\u00e8"}
{"text": "L\u2019exportation de l\u2019\u00e9nergie \u00e9olienne, une source de profits boud\u00e9"}
{"text": "@user ah \u00e7a c'est s\u00fablime. J'ai fait du volont\u00e9ariat en \u00e9cologie.", "label": "1"}
{"text": "L\u2019exportation de l\u2019\u00e9nergie \u00e9olienne, une source de profits boud\u00e9"}
{"text": "En France aussi, la police / gendarmerie est au service d'une minorit\u00e9... http"}
{"text": "La lutte contre la mosa\u00efque du manioc aux Comores: La s\u00e9lection vari\u00e9"}
{"text": "Il nous sort \"J'vais pas allumer la lumi\u00e8re parce que je veux pas faire marcher"}
{"text": "Cuba: Mesures contre les effets du changement http #Cuba", "label": "0"}
{"text": "@user @user @user Je mets au d\u00e9fis ce d\u00e9put\u00e9 bobo \u00e9cologiste d'a"}

```

4.) processed_german.jsonl (Pre-Processed by manually filtering out any jsonl objects with a “1” value for the “label” key, which would be the ‘neutral’ rating)

```
{"text": "@user korrekt! Verstehe sowas nicht...", "label": "0"}
{"text": "Der Dubbletimepart von Julien war ja mal sowas von genial! :D (@user http", "label": "0"}
{"text": "#Instachallenge #Day16 #what #i #am #reading #Fratzensammler #Horror #Wattpad http", "label": "0"}
{"text": "@user ah. Hatte nur bis radolfzell geschaut wegen ticket. Aber so isses fast normal", "label": "0"}
{"text": "@user \u00dcberall lauert Gefahr. Unverhofft. #Achtung", "label": "0"}
{"text": "Sonntag \u2714 Fitness \u2714 Nichts tun\u2714", "label": "2"}
{"text": "@user da werden wir uns wohl nicht einig werden, bef\u00fcrchte ich!", "label": "0"}
{"text": "@user Toll. Sehr l\u00f6blich. Sieht echt klasse aus.", "label": "2"}
{"text": "\"Absurde Unwahrheiten\" -Hoene\u00df stocksauer auf den \"Stern\" http #Absurde #Na", "label": "0"}
{"text": "RT @user: Linda, perfeita, guapa, Wundersch\u00f6n, Bella, Beautiful !! http", "label": "0"}
{"text": "Dank @user spiele ich jetzt ohne Schei\u00df Candy Crush Saga ... es regt mich ma\u00f", "label": "0"}
{"text": "Andreas Sp\u00e4ck aus der Crowd sucht ein originelles Hochzeitgeschenk... ;- ) Wer ka", "label": "0"}
{"text": "Nicht nur gg B\u00fcrger, sondern auch gg eigene Mitarbeiter mobben die Beh\u00f6rden", "label": "0"}
{"text": "@user OMG Titanium! Unbelievable Jahmene", "label": "2"}
{"text": "Unn\u00f6tiger schei\u00df. Ernsthaft.", "label": "0"}
{"text": "Bei http l\u00f6ffduft zur zeit EminemSing For The Moment also schnell einschalten :)", "label": "0"}
{"text": "RT @user: liebe @user, dein Duschgel schmeckt nicht #tranzparenztweet", "label": "0"}
{"text": "@user Aw, na gut, dann schlaf fein, klein Delalein :)", "label": "2"}
{"text": "@user kann man viel gelesen haben davon, nona, aber grunds\u00e4tzlich so nat\u00fcr", "label": "0"}
{"text": "@user kale chips= life", "label": "2"}
{"text": "Arschl\u00f6cher!!! MOBIL http", "label": "0"}
{"text": "RT @user: Das beste kommt zum Schluss? Ils\u00f6: \"Manchmal kommt der Beste auch er", "label": "0"}
{"text": "RT @user: Wach liegen, Gedankenkarussell, unbeantwortete Fragen, Bangen, Hoffen, Zwe", "label": "0"}
{"text": "@user Mich haben se nach vorne geschickt, konnt ne stille stehen ^^", "label": "2"}
```

5.) processed_hindi.jsonl (Pre-Processed by manually filtering out any jsonl objects with a “1” value for the “label” key, which would be the ‘neutral’ rating)

```
{"text": "CONGRESS na ye party kabhi bani hoti na india ka partition hota nd na hi humari coun", "label": "0"}
{"text": "ha ha ha ha ha ha ha . . . :d ye mast tha . . . cute aur comic . . .", "label": "0"}
{"text": "kejariwal tum apana soch ka dayra badho ab is misunderstanding se kam nahi chalane wa", "label": "0"}
{"text": "jai hind", "label": "2"}
{"text": "result me konsa bc univrsty tope mari ... jo itna wait kr rha h ", "label": "0"}
{"text": "rt @mukhijanidhi: hai teri hasti aisi jo dekhne wo mar mite #msgyouthicon #msgrevolut", "label": "0"}
{"text": "Dhyan rakh ?? ", "label": "0"}
{"text": "aisa school ho to me kabhi ghar hi na jau . . . ", "label": "2"}
{"text": "pratiyogita darpan ki spelling mistake hai :)", "label": "0"}
{"text": "i'll never forget that first girl i was crazy about in 5th grade . i still got her v", "label": "0"}
{"text": "BC porn site ki tarah advertise karo tum bas ..... !!!! male hoe .. ", "label": "0"}
{"text": "hahaha sai m . . ek baar class se nikal jate the to pura school naap ke wapas ate t", "label": "0"}
{"text": "post of the night: anonymous (id: excv6jkr) 11/02/12(fri)00:52:39 no . 434329650 thi", "label": "0"}
{"text": "Jaanna main bol rahi hu ki,tum mere twits dekho :/", "label": "2"}
{"text": "Did ", "label": "0"}
{"text": "rt psharma2525: rt komalinsaan: gurmeetramrahim #lovetoseemsg #msgincinemas schi agr", "label": "0"}
{"text": "Woh bhi ", "label": "0"}
```

```

{"text": "rare , 300kg meteorite discovered in poland , biggest in eastern europe and may pr
{"text": "band karo ye atyachaar. #indvsuae", "label": "0"}
{"text": "durdarshan wah wah . . . . .", "label": "2"}
{"text": "Ruk ", "label": "0"}
{"text": "Salman bai ap kese hen ma b ap ka fen hun", "label": "2"}
{"text": "aur jab koi ek excuse maar deta tha toh dusra kehta saale ye maine socha hua tha tun
{"text": "kiranji ko delhi ka cm ummidwar ghosit karna modiji ki doodarshita wa samghdaari ko

```

6.) processed_italian.jsonl (Pre-Processed by manually filtering out any jsonl objects with a “1” value for the “label” key, which would be the ‘neutral’ rating)

```

{"text": "@user @user Ma Ferrero? il compagno Ferrero? ma il suo partito esiste ancora? allora
{"text": "Non vedi l'ora che venga qui almeno lo sentir\u00f2 pi\u00f9 spesso e potr\u00f2 fin
{"text": "@user @user solo che poi arriva @user e decide sempre tutto lui", "label": "0"}
{"text": "@user Tutta la pasticceria! \ud83d\udc8f Grazie amore mio! \u2764", "label": "2"}
{"text": "Quanto mi stai cazzo troia", "label": "0"}
{"text": "Perch\u00e9 alla fine siamo una famiglia e la famiglia non va mai abbandonata\ud83d\
{"text": "@user @user @user @user @user calcio \u00e8 la quintessenza della noia. Sport frenet
{"text": "@user Ciao!Mia sorella ha un canale YouTube che si chiama corinnelife97potresti visi
{"text": "cazzo sono tutti sti tweet in tl di \\\\\"innocente\\\\" \\\\\"colpevole\\\\" ao", "label"
{"text": "La bellezza della mie Steve Madden Fringly \ud83d\udc97 http", "label": "2"}
{"text": "@user per\u00f2 ha detto che \u00e8 collegato positivamente alla pedofilia quindi ci
{"text": "La pi\u00f9 grande prova d'amore \u00e8 sempre stata togliersi dai coglioni.", "label
{"text": "@user @user Quando invece riaprii i miei occhi mi accorsi che non era un sogno...Era
{"text": "\u2013 tomlinsik;Profilo molto carino in generale abbinato molto bene.", "label": "2"
{"text": "Io faccio di tutto e lei ha ancora la faccia di ignorarmi perch\u00e9 sono cos\u00ec
{"text": "@user @user @user @user @user @user prima ora e dopo \ud83d\ude0d\ud83d\ude0d\ud83d\
{"text": "@user quanto ti odio da uno a dieci", "label": "0"}
{"text": "@user @user @user ti vogliamo bene anche noi Harry", "label": "2"}
{"text": "Io mi lamento della gente che scrive ancora \\\\\"freddy mercury\\\\" ma anche quella ch
{"text": "Oggi cominciamo cos\u00ec! Con i nostri saluti per voi di venerd\u00ec scorso dopo \
{"text": "Dal prossimo anno Audi non sar\u00e0 pi\u00f9 uno sponsor del Milan il testimone pass
{"text": "IL MIO DIARIO - Pagine d'estate quando ti innamori piano piano dietro uno schermo. G
{"text": "@user @user il fatto che una persona che vada in discoteca implichi che questa si dro
{"text": "@user che la forza sia con noi", "label": "2"}

```

7.) processed_portuguese.jsonl (Pre-Processed by manually filtering out any jsonl objects with a “1” value for the “label” key, which would be the ‘neutral’ rating)

```

{"text": "todos os meus favoritos na prova de elimina\u00e7\u00e3o #MasterChefBR", "label": "0"
{"text": "Augusto Cury \u00e9 o cara\ud83d\ude09 entrevista myito boa do #ConversaComBial", "la
{"text": "Mano Vitor, t\u00e9 muito feio! Decep\u00e7\u00e3o. #MasterChefBR", "label": "0"}
{"text": "Morrendo de amores pelo @user no programa do @user. #TheNoite", "label": "2"}
{"text": "MIRIAN CHOORANDO PELA PROVA PASSADA #MasterChefBR", "label": "0"}
{"text": "a cada epis\u00f3dio q passa Fabrizio continua um gato ne \ud83d\ude0f #MasterChefBR
{"text": "Traz o @user e tira a Mirian Pvr #MasterChefBR", "label": "0"}
{"text": "A maior e melhor cantora que voc\u00ea respeita @user no @user. #SandyLeah #Encontro
{"text": "Imagina que insuport\u00e1vel ter de dar de comer pra uma gente que calcula CADA CAL

```

```

{"text": "Baixa caloria? Frango desfiado e salada. T\u00e9 \u00f3timo #MasterChefBR", "label": "2"}
{"text": "A\u00ed que agonia essa Marian, meu Deus do c\u00e9u, mulher!! #MasterChefBR", "label": "2"}
{"text": "\"Eu adoro fazer piquenique no TEMPRO\u201d haha te amo yuko #MasterChefBR", "label": "2"}
{"text": "\"uma mulher mais nova com um homem mais velho \u00e9 normal\" normal dizer que ela \u00e9 normal", "label": "2"}
{"text": "Que amor esse @user no #Encontro \u2764\u2764\u2764", "label": "2"}
{"text": "Vanessa da Mata acabou com a fofurice do parto da mala da Bela Gil kkkkkk #AltasHoras", "label": "2"}
{"text": "f\u00e9tima bernardes maravilhosa no #VideoShowAoVivo", "label": "2"}
{"text": "Al\u00f4 NET, voces t\u00e3o fora do ar?! Arrumem por favor que eu quero assistir #MasterChefBR", "label": "2"}
{"text": "#MasterChefBR Hummm deu vontade de comer uma sardinha com p\u00e3o kkk partiu al\u00f4", "label": "2"}
{"text": "Pq o programa n\u00e3o ta ao vivo? @user #MaisVoc\u00ea", "label": "0"}
{"text": "Mr Catra dando um show no The Noite ! Falando da pol\u00edtica no pa\u00eds \ud83d\udd25", "label": "2"}
{"text": "Se algu\u00e9m me fala que \u00e9 termog\u00eanico, eu devolvo o prato na hora #MasterChefBR", "label": "2"}
{"text": "Ai eu me derreto todinha com a Paola \ud83d\udd25 \ud83d\udd25 \ud83d\udd25 #MasterChefBR", "label": "2"}
{"text": "#AltasHoras assistindo o programa com uma cx de len\u00e7\u00f3is ao lado! Haja \u2764\u2764", "label": "2"}
{"text": "Legal est\u00e1 prova tomara que tenha mais #DomingoLegal", "label": "2"}

```

8.) processed_spanish.jsonl (Pre-Processed by manually filtering out any jsonl objects with a “1” value for the “label” key, which would be the ‘neutral’ rating)

```

{"text": "@user jajajaja dale, hacete la boluda vos jajaja igual a vos nunca se te puede tomar", "label": "2"}
{"text": "@user MAAAAE RAJADO! Pero lo bueno es q uno se va independizando!y logrando metas", "label": "2"}
{"text": "Bueno hoy fui a almorzar a Nanay con otras 3 dras xq la capacitaci\u00f3n mal organiz", "label": "2"}
{"text": "@user \u00a1Hola Tom\u00e9is! \u00bfHab\u00e9is visto los nuevos #dinos de #TierraMagica", "label": "2"}
{"text": "@user la hijueputa tela se me sali\u00f3. yo quer\u00eda volver a quedar acostada.", "label": "2"}
{"text": "@user pues no est\u00e1 nada mal", "label": "2"}
{"text": "@user quiz\u00e1 para profesionales no sea mucho,pero hay no remunerados principalmen", "label": "2"}
{"text": "Hora de seguir so\u00f1ando muy bonito", "label": "2"}
{"text": "Me estoy muriendo. Ojal\u00e1 mi jefa me haga trabajar viernes y s\u00e1bado en la ma", "label": "2"}
{"text": "La felicidad tiene un nombr #Tailandia Si es tu PRIMERA VEZ, apunta estos #consejos", "label": "2"}
{"text": "@user @user justamente ah\u00ed es tu equivocada suposici\u00f3n, qxq estudi\u00f3 en", "label": "2"}
{"text": "@user Alpha suerte y ten un feliz a\u00f1o 2017", "label": "2"}
{"text": "Que viene Nonpa y fijo es para mayores. Ya lloro.", "label": "0"}
{"text": "@user pues s\u00ed, s\u00ed lo que se siente. Pero pi\u00e9nselo como el \u00faltimo", "label": "2"}
{"text": "@user @user @user @user HAHAAHHA MMM NOS VAMOS EN 2 semanas y no tenemos ni hotel Pa", "label": "2"}
{"text": "@user Que bonito,y yo tengo una sorpresa para Ti ,que te gustara much\u00edsimo!", "label": "2"}
{"text": "No me sorprende que 2\u00e9 se haya separado, man. Pero a\u00fan as\u00ed me dio en el c", "label": "2"}
{"text": "@user pero yo te kiero", "label": "2"}
{"text": "@user tu te planteas comprar gr\u00e1fica. Yo si lo planteo tengo que comprar todo, m", "label": "2"}
{"text": "Cambiar \"El mejor sistema\u2122\" y la educaci\u00f3n m\u00e1s \u00e9tica en Espa\u00f1a se", "label": "2"}
{"text": "@user @user fallecido hace meses. Hay q contrastar.", "label": "0"}
{"text": "Que haya una planta de tamarindo afuera de la que posiblemente va a ser la casa en la", "label": "0"}
{"text": "no me deis confianza que luego os digo las cosas muy claras y os jode", "label": "0"}
{"text": "@user ay, gracias. Lo malo es que ese es solo para primer viaje. Pero mil gracias ig", "label": "0"}

```

2.4 A4 - Formatted JSON->Python Data Loading from Post-Classified Sentiment Ratings from ChatGPT & LLaMA 2

2.4.1 (full_tweety_mangos.py)

```
[ ]: parsed_arabic = [
    {
        "tweet": "          )      (          http vía␣
↪@user",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "          ␣
↪      ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "          http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "@user          \"          \",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": " -          :\" ... http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "          !      _ # @user _ # @user␣
↪ - # @user _ #",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "          - - - - #
↪      ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    }
]
```

```

    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "                _ # @user                                □
↪http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "                _ # _ #                                □
↪2017 # _ _ #",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "                # # # # # # # # # # # # # # □
↪ _ # # @user... http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "#star_news@Cheb_Khaled_                                □
↪ http ",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "                _ # http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "                @user #nadinenassibnjeim #nnn□
↪#star #actress... http",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user -                _ # http",

```

```

    "actual-sentiment": "2",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "@user
    ↪      .",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user - \      \      \      _ # _ # http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": " - : \      \      ↪
    ↪      . Part _ # 2",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user - _ # _ _ # http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": " : #",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "I liked a @user video from @user http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": " - _ # _ _ # http",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  }
}

```



```

    },
    {
        "tweet": "                ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "                ) (                "
        ↪ \ " ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "                ... http",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    }
]

parsed_english = [
    {
        "tweet": "Trying to have a conversation with my dad about vegetarianism,
        ↪ is the most pointless infuriating thing ever #caveman ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user You are a stand up guy and a Gentleman Vice President,
        ↪ Pence ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "@user @user @user Looks like Flynn isn't too pleased with me,
        ↪ he blocked me. You blocked by Flynn too @user ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "i'm not even catholic, but pope francis is my dude. like i
        ↪ just need him to hug me and tell me everything is okay. ",

```

```

        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "@user for al the crying you do about how middle America is_
↳left out-they have twice as much voting power ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Samsung to Bring Android 7.0 Nougat to Galaxy S6, S6 edge,_
↳Note 5, and Tab S2 - Softpedia News ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Israel deems comatose Gaza man who needs treatment in West_
↳Bank a security threat. #Palestine via @user ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "I will go so far to say s1 of westworld isn't just good, it's_
↳brilliant. A story within a story within a story about storytelling ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Iraqi Forces set to storm 3 areas of #Mosul, #AlQahira,_
↳#alMasarif & #alAmn#MosulOps#mosuloffensive#iraq#ISIS... ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "#NationalFastFoodDay Would love to live there. Chick-fil-A _
↳",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {

```

```

{
  "tweet": "Electoral College must reject Trump #rejecttrump_
↪#notmypresident #takingbackdemocracy #blacklivesmatter #nevertrump ",
  "actual-sentiment": "0",
  "chatgpt-sentiment": "0",
  "llama2-sentiment": "0"
},
{
  "tweet": "The Reputation Doctor weighs in on Tony Romo #NFL @user joins_
↪@user on #TheMorningRush LISTEN: ",
  "actual-sentiment": "2",
  "chatgpt-sentiment": "2",
  "llama2-sentiment": "0"
},
{
  "tweet": "Stop #fracking #Cuadrilla persecuting land defenders through_
↪the legal system #BankruptCuadrilla #ECOCIDE ",
  "actual-sentiment": "0",
  "chatgpt-sentiment": "0",
  "llama2-sentiment": "2"
},
{
  "tweet": "So proud of way @user & #trumpTransitionteam are molding_
↪strong leadership group for #America #TeamTrump #MakeAmericaGreatAgain ",
  "actual-sentiment": "2",
  "chatgpt-sentiment": "2",
  "llama2-sentiment": "2"
},
{
  "tweet": "@user Everyone's thinking far too short-termed. When all_
↪fossil fuels run out with no substitutes, then the crisis will come. ",
  "actual-sentiment": "0",
  "chatgpt-sentiment": "2",
  "llama2-sentiment": "0"
},
{
  "tweet": "@user @user - #ScreamQueens so lucky to get to work with TK._
↪",
  "actual-sentiment": "2",
  "chatgpt-sentiment": "2",
  "llama2-sentiment": "2"
},
{
  "tweet": "@user @user @user Why is it that today society is casually_
↪comfortable about being publicly distasteful?#deplorables ",
  "actual-sentiment": "0",

```

```

        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "This is a big deal and a smart move by Microsoft: ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Ben Carson for Housing & Urban Development??  I just can't  ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Digesting while watching #ScreamQueens s1 Thanksgiving &
        Black Friday eps. SO SO GOOD!! ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Israel's New Racism: Persecution of African Migrants in Holy
        Land #gaza #palestine #israel #BDS ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "I've been listening to Leonard Cohen nonstop for the past
        month. Only to find out he passed literally weeks ago. Wow my heart is so
        overwhel ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user She just didn't get them in areas were she needed them.
        Lots of voter suppression going on. Hacking & tampering ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Listen to #NBAAwards Winner @user interview on @user ",

```

```

        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    }
]

parsed_french = [
    {
        "tweet": "Royal: le président n'aime pas les pauvres? \"c'est n'importe_
↪quoi\" http ...",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "#press #seo: Nager vert avec les piscines écologiques http_
↪#piscineécologique",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "#Biodiversité #Gard : Des Lézards ocellés détruits sans_
↪autorisation à Saint-Gilles ! http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Etat des lieux et analyse des mécanismes de soutien public_
↪aux énergies renouvelables en France http",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Les espèces disparaissent 1.000 fois plus depuis que l'Homme_
↪existe | @user http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Une entreprise allemande propose une solution éco-efficace de_
↪ferme urbaine http #NRJ(Canada) #RegionsofFrance",
        "actual-sentiment": "2",

```

```

        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Réchauffement : l'ONU prédit un véritable «enfer climatique»  

↳d'ici 2050 http http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user @user Ah Oui, les éoliennes, pourrait aussi les  

↳éloigner .. ouf Merci...J'en veux une dans ma cour...",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Info: Éoliennes - Le résultat de l'enquête publique est tombé  

↳pour Esnes et Montzéville : avis défavorable",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Réchauffement climatique: il est grand temps d'agir - http",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "#GCO : les écologistes strasbourgeois sont \"stupéfaits\" -  

↳http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "J'aime ma bio parce que je mange bio et je travail dans la  

↳biologie et je roule en biocarburant.",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Polémique autour de l'abattage massif de bouquetins en  

↳Haute-Savoie | Eco(lo) http",

```

```

    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Echange autour du thème de l'agriculture raisonnée à la_
↪Réunion http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Un réseau de journalistes pour lutter contre les_
↪négaționnistes du changement climatique. http",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user \"Illimitées\", cela fera sourire les_
↪environnementalistes. Et le thème du déclin américain est récurrent depuis_
↪les années 1970...",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "L'exportation de l'énergie éolienne, une source de profits_
↪boudée par Hydro-Québec | Michel Lagacé... http",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user ah ça c'est sûr. J'ai fais du volontariat en écologie.",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "L'exportation de l'énergie éolienne, une source de profits_
↪boudée par Hydro-Québec | Michel Lagacé http",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {

```



```

        "tweet": "En France aussi, la police / gendarmerie est au service d'une
↳minorité... http #Sivens #Écologie http",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "La lutte contre la mosaïque du manioc aux Comores: La
↳sélection variétale pour lutter contre la mosaïque du manioc http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Il nous sort \"J'avais pas allumer la lumière parce que je
↳veux pas faire marcher des centrales électriques.\" Ah.",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Cuba: Mesures contre les effets du changement http #Cuba",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user @user @user Je mets au défis ce député bobo écologiste
↳d'aller au chantier de Saint-Nazaire #chiche",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "2"
    }
]

parsed_german = [
    {
        "tweet": "@user korrekt! Verstehe sowas nicht...",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Der Dubbletimepart von Julien war ja mal sowas von genial! :D
↳(@user http",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",

```

```

        "llama2-sentiment": "2"
    },
    {
        "tweet": "#Instachallenge #Day16 #what #i #am #reading #Fratzensammler_
↪#Horror #Wattpad http",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user ah. Hatte nur bis radolfzell geschaut wegen ticket.
↪Aber so isses fast normal ;)",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user Überall lauert Gefahr. Unverhofft. #Achtung",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Sonntag Fitness Nichts tun ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user da werden wir uns wohl nicht einig werden, befürchte
↪ich!",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user Toll. Sehr löblich. Sieht echt klasse aus.",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "\"Absurde Unwahrheiten\" -Hoeneß stocksauer auf den \"Stern\"_
↪http #Absurde #Nachrichten #N24",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    }

```

```

    },
    {
      "tweet": "RT @user: Linda, perfeita, guapa, Wunderschön, Bella,␣
↳Beautiful !! http",
      "actual-sentiment": "2",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "2"
    },
    {
      "tweet": "Dank @user spiele ich jetzt ohne Scheiß Candy Crush Saga ...␣
↳es regt mich maßlos auf aber aufhören kann ich auch ned -.-",
      "actual-sentiment": "0",
      "chatgpt-sentiment": "0",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "Andreas Späck aus der Crowd sucht ein originelles␣
↳Hochzeitgeschenk... ;-) Wer kann helfen? Merci ! Chuy You... http",
      "actual-sentiment": "2",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "1"
    },
    {
      "tweet": "Nicht nur gg Bürger, sondern auch gg eigene Mitarbeiter␣
↳mobben die Behörden. Diese Fälle bearbeitet http",
      "actual-sentiment": "0",
      "chatgpt-sentiment": "0",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "@user OMG Titanium! Unbelievable Jahmene",
      "actual-sentiment": "2",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "2"
    },
    {
      "tweet": "Unnötiger scheiß. Ernsthaft.",
      "actual-sentiment": "0",
      "chatgpt-sentiment": "0",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "Bei http l uft zur zeit EminemSing For The Moment also␣
↳schnell einschalten :)",
      "actual-sentiment": "2",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "2"
    }
  ]

```

```

    },
    {
      "tweet": "RT @user: liebe @user, dein Duschgel schmeckt nicht_
↳#tranzparenztweet",
      "actual-sentiment": "0",
      "chatgpt-sentiment": "0",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "@user Aw, na gut, dann schlaf fein, klein Delalein :)",
      "actual-sentiment": "2",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "@user kann man viel gelesen haben davon, nona, aber_
↳grundsätzlich so natürlich lächerlich, die Liste! ff",
      "actual-sentiment": "0",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "@user kale chips= life",
      "actual-sentiment": "2",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "2"
    },
    {
      "tweet": "Arschlöcher!!! MOBIL http",
      "actual-sentiment": "0",
      "chatgpt-sentiment": "0",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "RT @user: Das beste kommt zum Schluss? Ilsö: \"Manchmal kommt_
↳der Beste auch erst in den letzten 45min.\"... http",
      "actual-sentiment": "2",
      "chatgpt-sentiment": "2",
      "llama2-sentiment": "0"
    },
    {
      "tweet": "RT @user: Wach liegen, Gedankenkarussell, unbeantwortete_
↳Fragen, Bangen, Hoffen, Zweifeln, Stolz, Verblendung, Leere, Einsamkeit...",
      "actual-sentiment": "0",
      "chatgpt-sentiment": "0",
      "llama2-sentiment": "0"
    },
  },

```

```

    {
        "tweet": "@user Mich haben se nach vorne geschickt, konnt ne stille_
↳ stehen ^^",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    }
]

parsed_hindi = [
    {
        "tweet": "CONGRESS na ye party kabhi bani hoti na india ka partition_
↳ hota nd na hi humari country itni khokhli hoti @ ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "ha ha ha ha ha ha ha . . . . :d ye mast tha . . . ._
↳ cute aur comic . . . . :)",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "kejariwal tum apana soch ka dayra badho ab is_
↳ misunderstanding se kam nahi chalane wala hai tumhara",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "jai hind",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "result me konsa bc univrsty tope mari ... jo itna wait kr rha_
↳ h ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "rt @mukhijanidhi: hai teri hasti aisi jo dekhe wo mar mite_
↳ #msgyouthicon #msgrevolution",

```

```

        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Dhyan rakh ?? ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "aisa school ho to me kabhi ghar hi na jau . . . ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "pratiyogita darpan ki spelling mistake hai :)",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "i'll never forget that first girl i was crazy about in 5th↵
↵grade . i still got her valentine day card in my secret safe . ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "BC porn site ki tarah advertise karo tum bas ..... !!!! male↵
↵hoe .. ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "hahaha sai m . . ek baar class se nikal jate the to pura↵
↵school naap ke wapas ate the",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "post of the night: anonymous (id: excv6jkr) 11/02/12(fri)00:
↵52:39 no . 434329650 this thread is like watching a monkey fuck a coconut",
        "actual-sentiment": "0",

```

```

    "chatgpt-sentiment": "0",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "Jaanma main bol rahi hu ki,tum mere twits dekho :/",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Did ",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "rt psharma2525: rt komalinsaan: gurmeetramrahim #lovetoseemsg_
    ↪#msgincinemas schi agr duniya ye movi dekh le to yaha swrg bn skta h",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "Woh bhi ",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "rare , 300kg meteorite discovered in poland , biggest in_
    ↪eastern europe and may provide clues about earth's core: http/URL",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "band karo ye atyachaar. #indvsuae",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "durdarshan wah wah . . . . ",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  },

```

```

{
  "tweet": "Ruk ",
  "actual-sentiment": "0",
  "chatgpt-sentiment": "2",
  "llama2-sentiment": "0"
},
{
  "tweet": "Salman bai ap kese hen ma b ap ka fen hun",
  "actual-sentiment": "2",
  "chatgpt-sentiment": "2",
  "llama2-sentiment": "0"
},
{
  "tweet": "aur jab koi ek excuse maar deta tha toh dusra kehta saale ye_
↪maine socha hua tha tune kyun bol diya",
  "actual-sentiment": "0",
  "chatgpt-sentiment": "0",
  "llama2-sentiment": "0"
},
{
  "tweet": "kiranji ko delhi ka cm ummidwar ghosit karna modiji ki_
↪doodarshita wa samghdaari ko darshata hai . ",
  "actual-sentiment": "2",
  "chatgpt-sentiment": "0",
  "llama2-sentiment": "2"
}
]

parsed_italian = [
  {
    "tweet": "@user @user Ma Ferrero? il compagno Ferrero? ma il suo_
↪partito esiste ancora? allora stiamo proprio frecati !!!",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Non vedi l'ora che venga qui almeno lo sentirò più spesso e_
↪potrò finalmente stare con lui alla faccia di chi mi ha friendzonato.",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user @user solo che poi arriva @user e decide sempre tutto_
↪lui",
    "actual-sentiment": "0",

```



```

    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user Tutta la pasticceria!  Grazie amore mio! ",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "Quanto mi stai cazzo troia",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Perché alla fine siamo una famiglia e la famiglia non va mai_
↪ abbandonata #WeWillMeetAgain1D http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "@user @user @user @user @user calcio è la quintessenza della_
↪ noia. Sport frenetici come basket o volley no",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user Ciao!Mia sorella ha un canale YouTube che si chiama_
↪ corinnelife97potresti visitarlo e iscriverti se ti piace?Grazie ",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "cazzo sono tutti sti tweet in tl di \\\\\"innocente\\\\\\"_
↪ \\\\\"colpevole\\\\\\" ao",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "La bellezza della mie Steve Madden Fringly  http",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "0",

```

```

        "llama2-sentiment": "2"
    },
    {
        "tweet": "@user però ha detto che è collegato positivamente alla
↪pedofilia quindi ci vede qualcosa di buono",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "La più grande prova d'amore è sempre stata togliersi dai
↪coglioni.",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user @user Quando invece riaprii i miei occhi mi accorsi che
↪non era un sogno...Era tutto nitido e reale!Grigio e freddo...",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "- tomlinsik;Profilo molto carino in generale abbinato molto
↪bene.",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Io faccio di tutto e lei ha ancora la faccia di ignorarmi
↪perché sono così stupida? L'avrei dovuta mandare a fanculo da tempo.",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user @user @user @user @user @user prima ora e dopo ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "@user quanto ti odio da uno a dieci",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
    }

```

```

        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user @user @user ti vogliamo bene anche noi Harry",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Io mi lamento della gente che scrive ancora \\\\\"freddy_
↪mercury\\\\\\" ma anche quella che scrive \\\\\"jhonny cash\\\\\\" non scherza",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Oggi cominciamo così! Con i nostri saluti per voi di venerdì_
↪scorso dopo \\\\\"Notre Dame de Paris\\\\\\". http",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Dal prossimo anno Audi non sarà più uno sponsor del Milan il_
↪testimone passerà alla Piaggio...e via con gli Apetti!!",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "IL MIO DIARIO - Pagine d'estate quando ti innamori piano_
↪piano dietro uno schermo. Grazie Greta. @user http",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "@user @user il fatto che una persona che vada in discoteca_
↪implichi che questa si droghi è una fesseria...",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user che la forza sia con noi",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",

```

```

        "llama2-sentiment": "2"
    }
]

parsed_portuguese = [
    {
        "tweet": "todos os meus favoritos na prova de eliminação #MasterChefBR",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Augusto Cury é o cara entrevista myito boa do
↪#ConversaComBial",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Mano Vitor, tá muito feio! Decepção. #MasterChefBR",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Morrendo de amores pelo @user no programa do @user.
↪#TheNoite",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "MIRIAN CHOORANDO PELA PROVA PASSADA #MasterChefBR",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "a cada episódio q passa Fabrizio continua um gato ne
↪#MasterChefBR",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Traz o @user e tira a Mirian Pvr #MasterChefBR",
        "actual-sentiment": "0",

```

```

        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "A maior e melhor cantora que você respeita @user no @user.
↪#SandyLeah #Encontro #SandynoEncontro",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Imagina que insuportável ter de dar de comer pra uma gente
↪que calcula CADA CALORIA que come? Jesus... #MasterChefBR",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Baixa caloria? Frango desfiado e salada. Tá ótimo
↪#MasterChefBR",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Aí que agonia essa Marian, meu Deus do céu, mulher!!
↪#MasterChefBR",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "\"Eu adoro fazer piquenique no TEMPRO\" haha te amo yuko
↪#MasterChefBR",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "\"uma mulher mais nova com um homem mais velho é normal\"
↪normal dizer que ela ta com ele por dinheiro né? #Encontro",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "0",
        "llama2-sentiment": "0"
    },
    {

```

```

        "tweet": "Que amor esse @user no #Encontro  ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Vanessa da Mata acabou com a fofurice do parto da mala da_
↳Bela Gil kkkkkk #AltasHoras",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "fátima bernardes maravilhosa no #VideoShowAoVivo",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Alô NET, voces tão fora do ar?! Arrumem por favor que eu_
↳quero assistir #MasterChefBR hojeeee",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "#MasterChefBR Hummm deu vontade de comer uma sardinha com_
↳pãozinho kkk partiu abrir uma latinha! ",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Pq o programa não ta ao vivo? @user  #MaisVocê",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Mr Catra dando um show no The Noite ! Falando da política no_
↳país  #TheNoite @user @user",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {

```

```

        "tweet": "Se alguém me fala que é termogênico, eu devolvo o prato na_
↪hora #MasterChefBR",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Ai eu me derreto todinha com a Paola #MasterChefBR",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "#AltasHoras assistindo o programa com uma cx de lenços ao_
↪lado! Haja ",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "Legal está prova tomara que tenha mais #DomingoLegal",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    }
]

parsed_spanish = [
    {
        "tweet": "@user jajajaja dale, hacete la boluda vos jajaja igual a vos_
↪nunca se te puede tomar en serio te mando un abrazo desde Perú!",
        "actual-sentiment": "0",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "0"
    },
    {
        "tweet": "@user MAAAAE RAJADO! Pero lo bueno es q uno se va_
↪independizando!y logrando metas",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    },
    {
        "tweet": "Bueno hoy fui a almorzar a Nanay con otras 3 dras xq la_
↪capacitación mal organizada no nos dió almuerzo y encima nos mandan a comer_
↪2pm",
        "actual-sentiment": "0",

```

```

    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user ¡Hola Tomás! ¿Habéis visto los nuevos #dinos de
↪#TierraMagna? Es normal que haya colas antes de que comience el espectáculo",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user la hijueputa tela se me salió. yo quería volver a
↪quedar acostada.",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user pues no está nada mal",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "@user quizá para profesionales no sea mucho,pero hay no
↪remunerados principalmente femenino para quienes es un sueño, pasa en mi
↪país",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Hora de seguir soñando muy bonito",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Me estoy muriendo. Ojalá mi jefa me haga trabajar viernes y
↪sábado en la mañana. Porque cerrar va a ser fatal",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "La felicidad tiene un nombr #Tailandia Si es tu PRIMERA VEZ,
↪apunta estos #consejos! http http",

```



```

    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "@user @user justamente ahí es tu equivocada suposición, qxq_
    ↳estudió en 1 colegio ficho será pulcro? La religión le importa mda",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user Alpha suerte y ten un feliz año 2017",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "Que viene Nonpa y fijo es para mayores. Ya lloro.",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user pues sí, sé lo que se siente. Pero piénselo como el_
    ↳último esfuerzo, que ya ahorita llegan las vacaciones",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user @user @user @user HAHAAHHA MMM NOS VAMOS EN 2 semanas_
    ↳y no tenemos ni hotel Pa la primera noche  ",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "@user Que bonito,y yo tengo una sorpresa para Ti ,que te_
    ↳gustara muchísimo!",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "No me sorprende que 2ne1 se haya separado, man. Pero aún así_
    ↳me dio en el cora bc es un grupo legendario y empecé mi era kpop con ellas",

```

```

    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user pero yo te kiero",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user tu te planteas comprar gráfica. Yo si lo planteo tengo_
↪que comprar todo, mi pc es un venerable anciano",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Cambiar \"El mejor sistema\" y la educación médica en España_
↪será complicado, pero con gente como vosotros tengo fe @user @user",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "@user @user fallecido hace meses. Hay q contrastar.",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {
    "tweet": "Que haya una planta de tamarindo afuera de la que_
↪posiblemente va a ser la casa en la que vivas de vez en cuando, es una buena_
↪señal",
    "actual-sentiment": "2",
    "chatgpt-sentiment": "2",
    "llama2-sentiment": "2"
  },
  {
    "tweet": "no me deis confianza que luego os digo las cosas muy claras y_
↪os jode",
    "actual-sentiment": "0",
    "chatgpt-sentiment": "0",
    "llama2-sentiment": "0"
  },
  {

```

```

        "tweet": "@user ay, gracias. Lo malo es que ese es solo para primer_
↪viaje. Pero mil gracias igualmente",
        "actual-sentiment": "2",
        "chatgpt-sentiment": "2",
        "llama2-sentiment": "2"
    }
]

```

2.4.2 Extra Stuff, pt. 2.

Commented-Out Code from A3 (BERT & GPT-2 Pretrained Language Model-based Text-Classification Performance, Metrics, & Results)

```

##### CS410: Natural Language Processing, Fall 2023 - 11/13/2023
##### A3: Pretrained Language Models (PLMs), Dan Jang - Main Implementation
#### Objective: Exploring Natural Language Processing (NLP), by building a text-classifier
#### for a text classification task, predicting whether a piece of text is "positive" or "nega
#### ...focusing on two (2) pretrained language models (PLMs),
#### ...BERT (Bidirectional Encoder Representations from Transformers) & OpenAI's GPT-2 (Gener
#### ...and using the respective tokenizers to each PLM to perform the text-classification task

### 1.1.a) Logistic Regression algorithm using sklearn.linear_model.LogisticRegression
### https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.
### Returns four (4) things:
# I.) accuracy_score,
# II.) f1_score,
# III.) confusion_matrix,
# & IV.) classification_report.
def plm(model, xtrain, ytrain, xtest, ytest):

    lreg = LogisticRegression()

    lreg.fit(xtrain, ytrain)
    predictionresults = lreg.predict(xtest)

    return accuracy_score(ytest, predictionresults), f1_score(ytest, predictionresults), confu

### A3-Specific: Implementation functions for implementing the two pretrained language models,
## Pretrained Language Model (PLM) Tokenizer Implementation Function
def plmodel(words, model, tokenizer):
    if tokenizer.pad_token is None:
        raise ValueError("[Debug A3.1 - PLModel()]: Tokenizer has no padding token for the curr
    wordlist = tokenizer(words, return_tensors="pt", truncation=True, padding=True)

    with torch.no_grad():
        results = model(**wordlist)

    return results.last_hidden_state.mean(dim=1).squeeze().numpy()

```

```

# wordlist = words.split()
# vecs = [model[w] for w in wordlist if w in model]
# if vecs:
#     return sum(vecs) / len(vecs)
# else:
#     return [0] * model.vector_size

# ## GPT(transformer)-2 Model Implementation Function
# def bumblebee(words, model):
#     beds = [avgwordvec(w, model) for w in words]
#     return beds

def main(): #trainfile, testfile):
    print("Welcome, this is the main program for A3: Pretrained Language Models.")
    print("Written by Dan J. for CS410: Natural Language Processing, Fall 2023.")
    print("\nWe will use two (2) pretrained language models (PLM), BERT & GPT-2.\n...to create

# 1.0.I.A) Debug Statements #1a for dataset loading times:
print("\nLoading the training & testing datasets...")
# with open(trainfile, "r") as trainfile:
with open("sentiment_train.json", "r") as trainfile:
    #traindata = json.load(trainfile)
    for row in trainfile:
        traindata.append(json.loads(row))

trainframe = pandas.DataFrame(traindata)

# with open(testfile, "r") as testfile:
with open("sentiment_test.json", "r") as testfile:
    #testdata = json.load(testfile)
    for row in testfile:
        testdata.append(json.loads(row))

testframe = pandas.DataFrame(testdata)

# 1.0.I.B) Debug Statements #1b for dataset loading times:
print("Successfully loaded the training & testing datasets!\n")

## 1.0.1.) Initial Preprocessing of the training & testing data
## First, we isolate our two (2) columns, "review_title" & "stars."
## Second, we will convert values in the "stars" column so that 1 [negative] = 0 & 5 [positive]
## This will allow us to make the negative or positive sentiment a binary value-based thing.
trainframe = trainframe[['review_title', 'stars']]
trainframe['stars'] = trainframe['stars'].apply(lambda x: 1 if x == 5 else 0)

testframe = testframe[['review_title', 'stars']]
testframe['stars'] = testframe['stars'].apply(lambda x: 1 if x == 5 else 0)

```

```

## A3-Specific: From our Slack channel (#nlp_f23), using tip to only use 25% of training d
## Credits to Classmate Will McIntosh from the Slack thread started by classmate Saurav Ku
## & also, full credits to classmate Will McIntosh for the following code for GPU usage:

#### Credits to Will McIntosh (11/11/2023):
# Testing
print(f"Is a GPU available: {torch.cuda.is_available()}")
#print(f"Is this instance using a GPU?: {next(model.parameters()).is_cuda}")
#### From Slack, #nlp_f23.

trainframe = trainframe.sample(frac=1, random_state=69)
trainframe = trainframe.iloc[:int(0.25 * len(trainframe))]

# y2train = trainframe['stars']
# print("[A3 Debug Size-Print #3] y2train", len(y2train))

ytest = testframe['stars']
# print("[A3 Debug Size-Print #4] ytest", len(ytest))

## Evenly split frames
x3train1, x3train2 = train_test_split(trainframe, test_size=0.5, random_state=69)

y3train1 = x3train1['stars']
y3train2 = x3train2['stars']

#print("[A3 Debug Size-Print #1] x3train1 & x3train2", len(x3train1), len(x3train2))
## A3-Specific: Applying BERT & GPT-2 PLM Specific Tokenization
print("Tokenizing the training & testing datasets for BERT...")
x2train1 = x3train1['review_title'].apply(lambda x: plmodel(x, bertie, bertie_tokens))
#x2train1 = trainframe['review_title'].apply(lambda x: plmodel(x, bertie, bertie_tokens))
x2test1 = testframe['review_title'].apply(lambda x: plmodel(x, bertie, bertie_tokens))
print("BERT Tokenization has been applied to its training & testing datasets!")

#print("[A3 Debug Size-Print #2a] x2train1 & x2test1", len(x2train1), len(x2test1))

print("Tokenizing the training & testing datasets for GPT-2...")
x2train2 = x3train2['review_title'].apply(lambda x: plmodel(x, bumblebee, bumblebee_tokens))
#x2train2 = trainframe['review_title'].apply(lambda x: plmodel(x, bumblebee, bumblebee_tok
x2test2 = testframe['review_title'].apply(lambda x: plmodel(x, bumblebee, bumblebee_tokens))
print("GPT-2's tokenization has been applied to its training & testing datasets!")

#print("[A3 Debug Size-Print #2b] x2train2 & x2test2", len(x2train2), len(x2test2))

### 1.0.2b) Run Text-Classification Algorithms & Print the Model Results
print("-----\n")
print("Running text-classification model le training & testing datasets (with pretrained l

```

```

print("Running Logistic Regression algorithm, version A.) BERT...")
bed1accuracy, bed1f1, bed1cmatrix, bed1creport = plm(bertie, x2train1.tolist(), y3train1, )
print("..The data processing from our first PLM, BERT, is done!")

print("Running Logistic Regression algorithm, version B.) GPT-2...")
bed2accuracy, bed2f1, bed2cmatrix, bed2creport = plm(bumblebee, x2train2.tolist(), y3train2, )
print("..The data processing from our second PLM, GPT-2, is done!")

print("...All Done!")
print("-----\n")

print("Here are le results [Logistic Regression, with comparative results between two (2) PLMs, BERT & GPT-2]
print("Logistic Regression Algorithm, Version A: BERT Pretrained Language Model-based Text-Classification")
print("...Accuracy was found to be, ", bed1accuracy * percentness, "%,")
print("...F1 Score was found to be: ", bed1f1, ",")
print("...with a Confusion Matrix: \n", bed1cmatrix, ",")
print("...& lastly, the classification Report: \n", bed1creport)
print("-----\n")

print("Logistic Regression Algorithm, Version B: GPT-2 Pretrained Language Model-based Text-Classification")
print("...Accuracy was found to be, ", bed2accuracy * percentness, "%,")
print("...F1 Score was found to be: ", bed2f1, ",")
print("...with a Confusion Matrix: \n", bed2cmatrix, ",")
print("...& lastly, the classification Report: \n", bed2creport)
print("-----\n")

if __name__ == "__main__":
    main()

```

A3 Results: Raw Output from *BERT* & *OpenAI's GPT-2* Pretrained Language Models (PLMs), using the Logistic Regression Algorithm

Here are le results [Logistic Regression, with comparative results between two (2) PLMs, BERT & GPT-2]

Logistic Regression Algorithm, Version A: BERT Pretrained Language Model-based Text-Classification

...Accuracy was found to be, 91.55 %,
 ...F1 Score was found to be: 0.9158785465405676 ,
 ...with a Confusion Matrix:

```
[[911 89]
 [ 80 920]] ,
```

...& lastly, the classification Report:

	precision	recall	f1-score	support
0	0.92	0.91	0.92	1000
1	0.91	0.92	0.92	1000
accuracy			0.92	2000
macro avg	0.92	0.92	0.92	2000

weighted avg 0.92 0.92 0.92 2000

Logistic Regression Algorithm, Version B: GPT-2 Pretrained Language Model-based Text-Classification

...Accuracy was found to be, 87.4 %,

...F1 Score was found to be: 0.872598584428716 ,

...with a Confusion Matrix:

[[885 115]

[137 863]] ,

...& lastly, the classification Report:

	precision	recall	f1-score	support
0	0.87	0.89	0.88	1000
1	0.88	0.86	0.87	1000
accuracy			0.87	2000
macro avg	0.87	0.87	0.87	2000
weighted avg	0.87	0.87	0.87	2000

2.4.3 A2 Results: Raw Output from Word2Vec & GloVe Embedding Results, using the Logistic Regression Algorithm

Logistic Regression Algorithm, Version A: Word2Vec Pretrained Model-based Embeddings Performance

...Accuracy was found to be, 86.3 %,

...F1 Score was found to be: 0.86105476673428 ,

...with a Confusion Matrix:

[[1754 246]

[302 1698]] ,

...& lastly, the classification Report:

	precision	recall	f1-score	support
0	0.85	0.88	0.86	2000
1	0.87	0.85	0.86	2000
accuracy			0.86	4000
macro avg	0.86	0.86	0.86	4000
weighted avg	0.86	0.86	0.86	4000

Logistic Regression Algorithm, Version B: GloVe Pretrained Model-based Embeddings Performance,

...Accuracy was found to be, 69.69999999999999 %,

...F1 Score was found to be: 0.7313829787234042 ,

...with a Confusion Matrix:

[[1138 862]

```
[ 350 1650]] ,
...& lastly, the classification Report:
      precision    recall  f1-score   support

         0         0.76      0.57      0.65      2000
         1         0.66      0.82      0.73      2000

 accuracy          0.70      4000
macro avg          0.71      0.70      0.69      4000
weighted avg          0.71      0.70      0.69      4000
```

From A1 Results for Reference: Initial Full 80k-Row Processing Results Raw Output

Algorithm #1, Version A: Gaussian Naïve Bayes Performance, Metrics, & Results:

```
...Accuracy was found to be,  59.199999999999996 %,
...F1 Score was found to be:  0.3664596273291925 ,
...with a Confusion Matrix:
```

```
[[948  52]
 [764 236]] ,
```

```
...& lastly, the classification Report:
      precision    recall  f1-score   support

         0         0.55      0.95      0.70      1000
         1         0.82      0.24      0.37      1000

 accuracy          0.59      2000
macro avg          0.69      0.59      0.53      2000
weighted avg          0.69      0.59      0.53      2000
```

Algorithm #2, Version A: Logistic Regression Performance, Metrics, & Results:

```
...Accuracy was found to be,  92.7 %,
...F1 Score was found to be:  0.9272908366533865 ,
...with a Confusion Matrix:
```

```
[[923  77]
 [ 69 931]] ,
```

```
...& lastly, the classification Report:
      precision    recall  f1-score   support

         0         0.93      0.92      0.93      1000
         1         0.92      0.93      0.93      1000

 accuracy          0.93      2000
macro avg          0.93      0.93      0.93      2000
weighted avg          0.93      0.93      0.93      2000
```

Algorithm #1, Version B: Gaussian N  ive Bayes Performance, Metrics, & Results:

...Accuracy was found to be, 59.3 %,

...F1 Score was found to be: 0.36899224806201547 ,

...with a Confusion Matrix:

[[948 52]

[762 238]] ,

...& lastly, the classification Report:

	precision	recall	f1-score	support
0	0.55	0.95	0.70	1000
1	0.82	0.24	0.37	1000
accuracy			0.59	2000
macro avg	0.69	0.59	0.53	2000
weighted avg	0.69	0.59	0.53	2000

Algorithm #2, Version B: Logistic Regression Performance, Metrics, & Results:

...Accuracy was found to be, 92.80000000000001 %,

...F1 Score was found to be: 0.9281437125748503 ,

...with a Confusion Matrix:

[[926 74]

[70 930]] ,

...& lastly, the classification Report:

	precision	recall	f1-score	support
0	0.93	0.93	0.93	1000
1	0.93	0.93	0.93	1000
accuracy			0.93	2000
macro avg	0.93	0.93	0.93	2000
weighted avg	0.93	0.93	0.93	2000

[23]: ##### Jupyter Notebook -> PDF Conversion thingy

```
#!pip install nbconvert
```

```
!jupyter nbconvert a4-Multilingual-Large-Language-Models-dan-jang --to pdf
```

[NbConvertApp] Converting notebook a4-Multilingual-Large-Language-Models-dan-jang.ipynb to pdf

[NbConvertApp] Writing 378472 bytes to notebook.tex

```
[NbConvertApp] Building PDF
[NbConvertApp] Running xelatex 3 times: ['xelatex', 'notebook.tex', '-quiet']
[NbConvertApp] Running bibtex 1 time: ['bibtex', 'notebook']
[NbConvertApp] WARNING | b had problems, most likely because there were no
citations
[NbConvertApp] PDF successfully created
[NbConvertApp] Writing 526137 bytes to a4-Multilingual-Large-Language-Models-
dan-jang.pdf
```