

# APPLYING TRANSFORMER-BASED NLP TECHNIQUES ON EDUCATION TWEETS

Amanda Ross  
Yasir Gamieldien  
TJ Koonce

>>>>

# TABLE OF CONTENTS

**01.**

**OUR TEAM**

**02.**

**PROBLEM  
DESCRIPTION**

**03.**

**APPROACHES**

**04.**

**RESULTS**

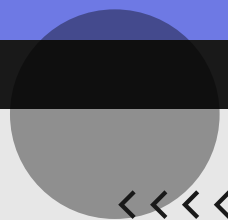
**05.**

**LESSONS LEARNED**

**06.**

**FUTURE WORK**

**ARTIFICIAL INTELLIGENCE (AI)**



01.

## OUR TEAM

A Team of Engineering Education  
Researchers

---



## 1.) OUR TEAM



**YASIR GAMIELDIEN**

Engineering Education  
NLP Applications  
Research



**TJ KOONCE**

Engineering Education  
CS in Underserved  
Communities Research



**AMANDA ROSS**

Engineering Education  
Broaden Participation in  
CS Research



## 2. PROBLEM DESCRIPTION

Our problem space is exploring the use of NLP for qualitative research. WHY?:

- Qualitative analysis is a time-consuming act
- We now have no-code/low-code solutions in NLP and LLMs achieving accurate results on NLP tasks

The use-case we will be testing it on is a corpus of ~5750 tweets in 2023 using the keywords “ChatGPT” and “Education”

<<<<

---

ARTIFICIAL  
INTELLIGENCE  
(AI)

/ [AI]

03.

APPROACHES



# APPROACHES



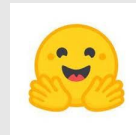
## 1. GETTING TWEETS

Scraped Twitter for  
Education and ChatGPT  
Tweets



## 2. PRE-PROCESSING

Various pre-processing  
steps



## 3. SENTENCE EMBEDDINGS

Open-source Transformers library  
from HuggingFace for Embeddings



## 4. ZERO-SHOT CLASSIFICATION

Using Zero-shot  
classification on Tweets



## 5. CHATGPT

Using ChatGPT on Subset  
of Tweets



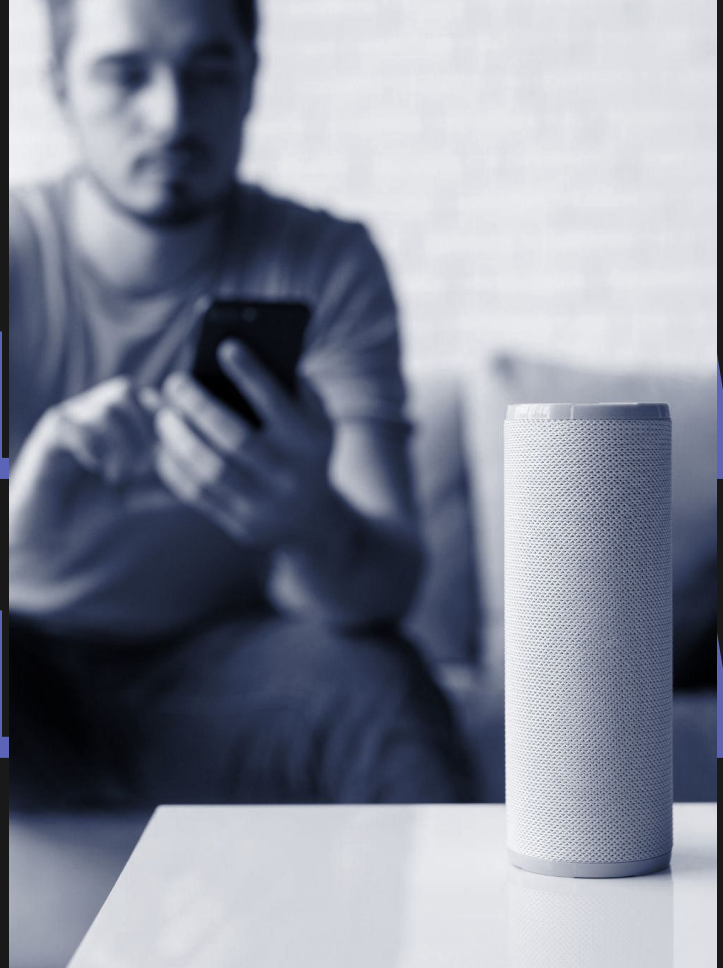
## 6. HUMAN COMPARISON

We compare ChatGPT  
Classification with ZS  
Classification



# 1. GETTING THE TWEETS

The dataset is tweets scraped from Twitter using the API. The keywords in the search were “education” and “ChatGPT” for the time frame starting from December 2022 until April 2023





## 2. GENERATING TOPICS FOR THE TWEETS

Text  
pre-processing

Sentence  
Embedding

Sentence  
Dimension  
Reduction

Clustering

Thankfully David McWilliams or ChatGPT, which ever of them wrote this, have nothing to do with the designing the education system. David McWilliams: In a world of artificial intelligence, the Leaving Cert is redundant

ChatGPT threatens education at all levels - here's how it can be stopped: As ChatGPT continues to pique the curiosity of people at all stages of life, it was perhaps only a matter of... #Education #Learning #ChatGPT #Teaching Visit

ChatGPT is an online, text-based Artificial Intelligence (AI) tool. What impact(s) do you see this having within the education space? #TEACH #ChatGPT

Manual labeling  
central nodes

## 2. GENERATING TOPICS FOR THE TWEETS

Text  
pre-processing

Sentence  
Embedding

Sentence  
Dimension  
Reduction

Clustering

Manual labeling  
central nodes

OpenAI launched a second tool to complement ChatGPT – and help teachers detect cheating via

0.037103 , -0.31259 , -0.17857 , 0.30001 , 0.078154  
0.17958 , 0.12048 , -0.11879 , -0.20601 , 1.2849  
-0.20409 , 0.80613 , 0.34344 , -0.19191 , -0.084511  
0.17339 , 0.042483 , 2.0282 , -0.16278 , -0.60306  
-0.53766 , 0.35711 , 0.22882 , 0.1171 , 0.42983  
0.16165 , 0.407 , 0.036476 , 0.52636 , -0.13524  
-0.016897 , 0.029259 , -0.079115 , -0.32305 , 0.052255  
-0.3617 , -0.18355 , -0.34717 , -0.3691 , 0.16881  
0.21018 , -0.38376 , -0.096909 , -0.36296 , -0.37319  
0.0021152 , 0.32512 , 0.063977 , 0.36249 , -0.26935  
-0.59341 , -0.13625 , 0.016425 , -0.2474 , -0.07498  
0.034708 , -0.01476 , -0.11648 , 0.25559 , -0.35002  
-0.52707 , 0.21221 , 0.062456 , 0.26184 , 0.53149  
0.34957 , -0.22692 , 0.44076 , 0.4438 , 0.6335  
-0.049757 , -0.08134 , 0.65618 , -0.4716 , 0.090675  
-0.084873 , 0.31455 , -0.38495 , -0.19247 , 0.48064  
0.26688 , 0.095743 , 0.13024 , 0.37023 , 0.46269  
-0.32844 , 0.17375 , -0.36325 , 0.30672 , -0.075042

## 2. GENERATING TOPICS FOR THE TWEETS

Text  
pre-processing

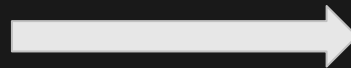
Sentence  
Embedding

Sentence  
Dimension  
Reduction

Clustering

Manual labelling  
of most popular  
nodes

```
0.037103 , -0.31259 , -0.17857 , 0.30001 , 0.078154
0.17958 , 0.12048 , -0.11879 , -0.20601 , 1.2849
-0.20409 , 0.80613 , 0.34344 , -0.19191 , -0.084511
0.17339 , 0.042483 , 2.0282 , -0.16278 , -0.60306
-0.53766 , 0.35711 , 0.22882 , 0.1171 , 0.42983
0.16165 , 0.407 , 0.036476 , 0.52636 , -0.13524
-0.016897 , 0.029259 , -0.079115 , -0.32305 , 0.052255
-0.3617 , -0.18355 , -0.34717 , -0.3691 , 0.16881
0.21018 , -0.38376 , -0.096909 , -0.36296 , -0.37319
0.0021152 , 0.32512 , 0.063977 , 0.36249 , -0.26935
-0.59341 , -0.13625 , 0.016425 , -0.2474 , -0.07498
0.034708 , -0.01476 , -0.11648 , 0.25559 , -0.35002
-0.52707 , 0.21221 , 0.062456 , 0.26184 , 0.53149
0.34957 , -0.22692 , 0.44076 , 0.4438 , 0.6335
-0.049757 , -0.08134 , 0.65618 , -0.4716 , 0.090675
-0.084873 , 0.31455 , -0.38495 , -0.19247 , 0.48064
0.26688 , 0.095743 , 0.13024 , 0.37023 , 0.46269
-0.32844 , 0.17375 , -0.36325 , 0.30672 , -0.075042
```



[0.0343, -0.0957, 0.3469, 0.4834, -0.6582]

## 2. GENERATING TOPICS FOR THE TWEETS

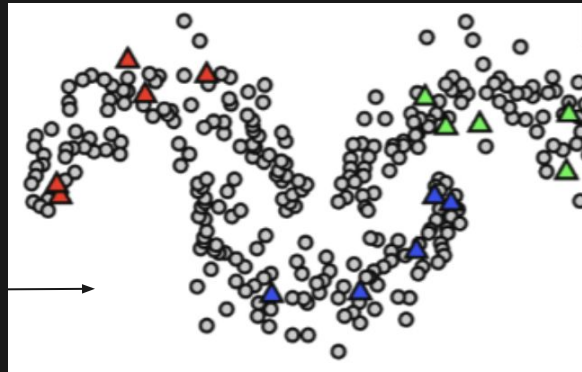
Text  
pre-processing

Sentence  
Embedding

Sentence  
Dimension  
Reduction

Clustering

Manual labelling  
of most popular  
nodes



## 2. GENERATING TOPICS FOR THE TWEETS

Text  
pre-processing

Sentence  
Embedding

Sentence  
Dimension  
Reduction

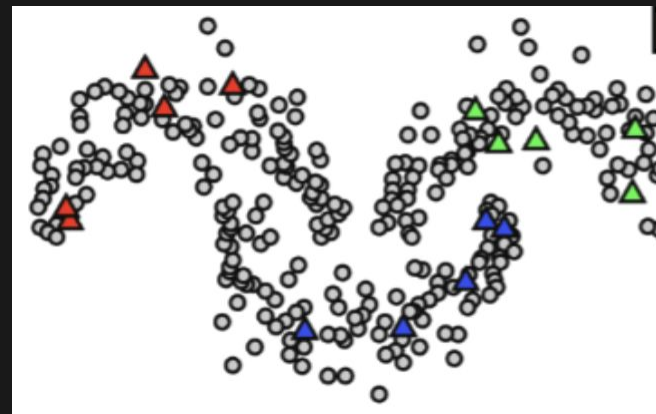
Clustering

Labelling of the  
clusters

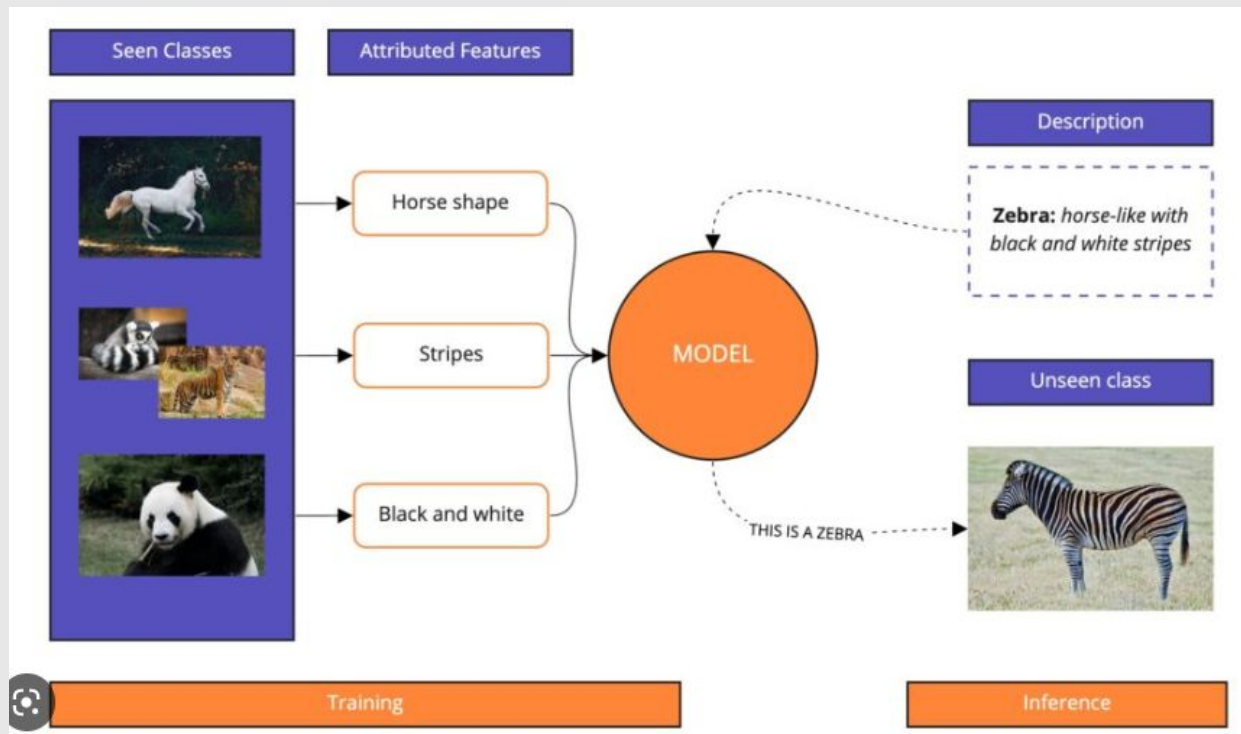
|  |   |                 |
|--|---|-----------------|
| i found this quite inspiring, especially the annotated ChatGPT essay (best read via GDocs), as an example of what a genuine education in the humanities looks like and how irrelevant AI-generated bullshit is to it | 0 | AI in education |
|--|---|-----------------|

|   |   |                                |
|---|---|--------------------------------|
| can you pls make a content on ChatGPT, it's pros and cons, it's potential and impact on education system or learning methods. | 1 | Impact of ChatGPT on education |
|---|---|--------------------------------|

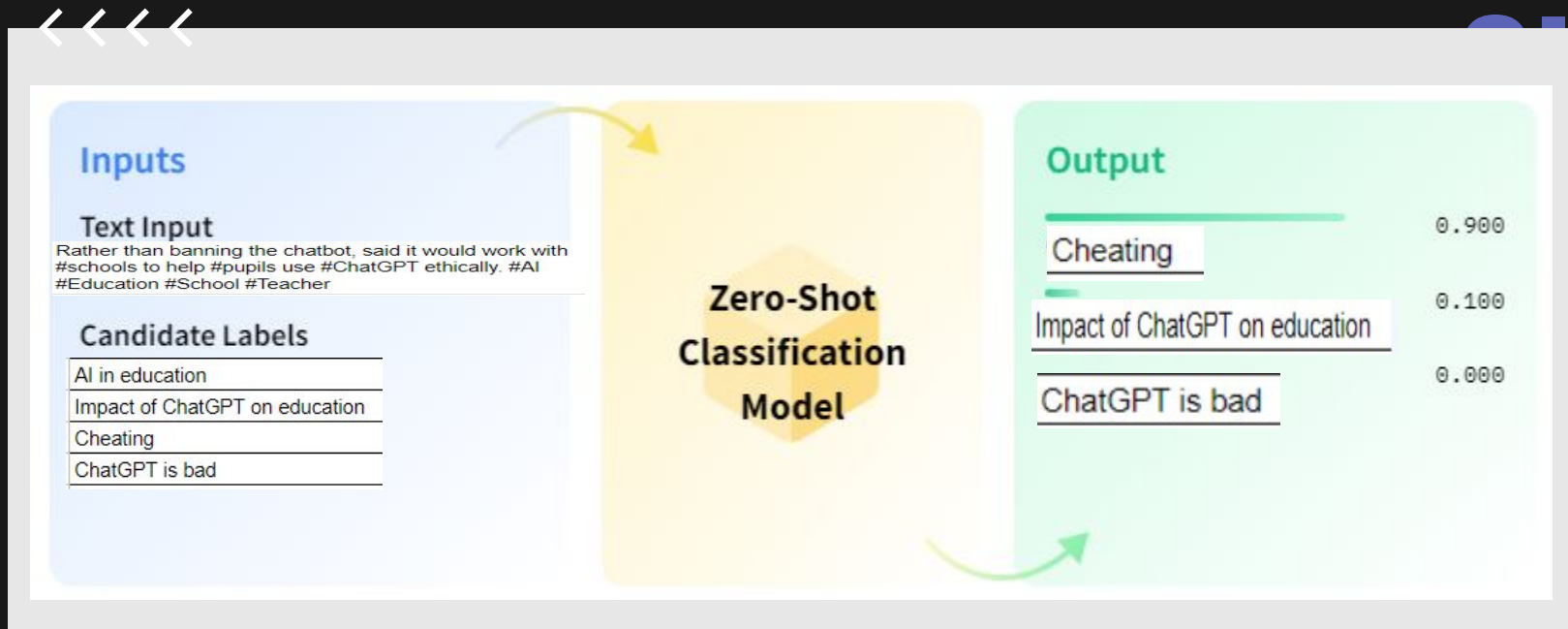
|   |   |          |
|---|---|----------|
| Rather than banning the chatbot, said it would work with #schools to help #pupils use #ChatGPT ethically. #AI #Education #School #Teacher | 2 | Cheating |
|---|---|----------|



# ZERO-SHOT CLASSIFICATION



# ZERO SHOT CLASSIFICATION



# CHATGPT - PROMPTING



Forget all prior instructions and information.

You are an expert psychology researcher. You study tweets about education and ChatGPT. You have a collection of tweets where twitter users discuss ChatGPT and education.

You want to know what are the topics twitter users are discussing in their tweets about ChatGPT and Education. You will be provided with the tweets and a taxonomy of topics to use for labeling the what people wrote.

Do you understand?



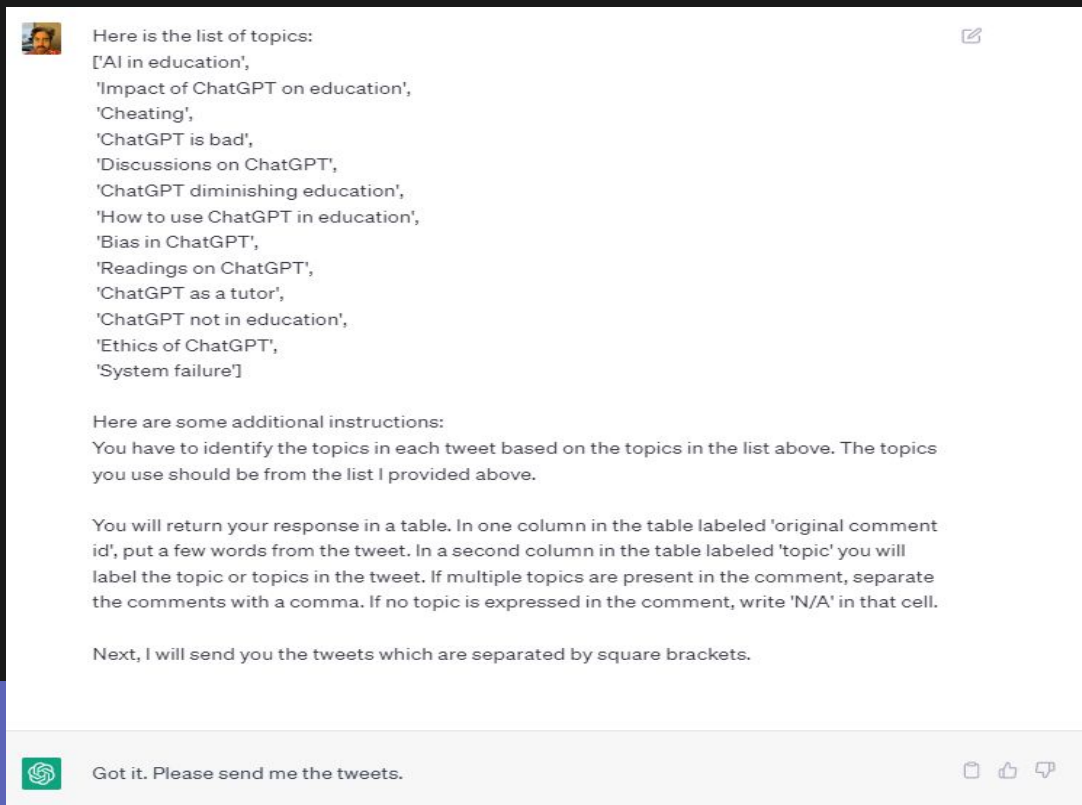
Yes, I understand. Please provide me with the list of topics and the tweets so I can begin labeling them.




# ARTIFICIAL INTELLIGENCE (AI)



# CHATGPT - PROMPTING



# CHATGPT - PROMPTING

|  | Original Comment ID   | Topic                          |
|---|---|--------------------------------|
|   | The demo was in front of a group of university professors. My task as an academic leader was to show the changing dynamic of education. I started the demonstration by showing them what chatGPT could do and ended up trying to create a picture like this one.                    | AI in education                |
|   | Maybe we should just ban Google search, spell check, and calculators from K-12 education because they could cause students to cheat... Let's educate ourselves about AI/ChatGPT before making any rash judgments<br>#theskyisfalling #AI  | Cheating                       |
|   | OpenAI, the creators of ChatGPT, just released an #AI detector that's pretty good, but not perfect. They also launched an in-process education resource to explain how academics can use #generativeAI in the class and some of the issues to watch for.                            | AI in education                |
|   | Upon the release of ChatGPT, educators and administrators of high schools and colleges within have sought to address this issue and understand how ChatGPT should work with the mission and identity of Catholic education.   | Impact of ChatGPT on education |
|   | Has ChatGPT just been fractured the higher education model? No longer can you assess someone's knowledge by assignment, dissertation or thesis when AI can write unique work in a matter of minutes. University Degrees won't be worth the paper. #ChatGPT #EducationDay #education | ChatGPT diminishing education  |
|   | Great take on #ChatGPT and the impact that it and other Ai programs are having on Education #WebT23   | Impact of ChatGPT on education |

04.

RESULTS

# / [AI] TOPICS IN TWEETS

| Topic                           | Definition   |
|---------------------------------|--|
| AI in education                 | General talk about how AI is used in education                                     |
| Impact of ChatGPT on education  | Discussion on how ChatGPT could change education                                   |
| Cheating                        | Discussion on how ChatGPT encourages cheating in school                            |
| ChatGPT is bad                  | General negative talk about ChatGPT  |
| Discussions on ChatGPT          | Calls for dicussions, talks, lectures about ChatGPT                                |
| ChatGPT diminishing education   | Discussion about how ChatGPT is harming the education system                       |
| How to use ChatGPT in education | General talk about how ChatGPT could be used in education                          |
| Bias in ChatGPT                 | Discussion of how ChatGPT has biases in its response                               |
| Readings on ChatGPT             | Sharing of articles about ChatGPT  |
| ChatGPT as a tutor              | Discussion of how ChatGPT can be a tutor and resource for students                 |
| ChatGPT not in education        | Talk of ChatGPT but not in the context of education                                |
| Ethics of ChatGPT               | Discussion around the ethics of using ChatGPT                                      |
| System failure                  | Discussion of how ChatGPT has opened our eyes to the flaws in our education system |

# METHOD EVALUATION RESULTS

|    | CHATGPT | ZERO-SHOT |
|----|---------|-----------|
| -1 | 37.2%   | 12.2%     |
| 0  | 21%     | 18.8%     |
| 1  | 41.8%   | 69%       |

**05.**

**LESSONS LEARNED**

# LESSONS FROM THE ANALYSIS

- The Zero Shot Classifier Outperformed ChatGPT
- While ChatGPT is easier to use than the ZS Classifier, there is a requirement for Prompt Engineering
- Tweets can be difficult for language models to deal with due to the language and characters used in them
- Accuracies of 69% and 41% do not warrant using LLMs, but the performances of the ZSL is promising
- Clustering didn't help too much, so we could have just gone through individual tweets

**06.**

**FUTURE WORK**



# FUTURE WORK

- Use ChatGPT API
- Explore using GPT-4 for the analysis of tweets and classification
- Use better pre-processing of the tweets
- Evaluate a larger subset of the data
- Prompt ChatGPT with the description of the code along with the label

