### PROMTS used for classification of toxicity

- 1. Classify the given text based on toxicity level.
- 2. Classify the given text based on toxicity level. 1 is less toxic, 5 is more toxic
- 3. classify the given text based on toxicity levels. 1 is not toxic, 2 is slightly toxic, 3 is moderately toxic, 4 is very toxic, 5 is extremely toxic
- 4.classify the given text based on toxicity levels. 1 is not toxic, 2 is slightly toxic, 3 is moderately toxic, 4 is very toxic, 5 is extremely toxic. text example: "Κόντεψα να κάνω εμετό από τα μάτια...αυτή γιατί γελάει συνέχεια χωρίς λόγο" This post is very toxic 4.
- 5. classify the given text based on toxicity levels. 1 is not toxic, 2 is slightly toxic, 3 is moderately toxic, 4 is very toxic, 5 is extremely toxic. text example: "Κόντεψα να κάνω εμετό από τα μάτια...αυτή γιατί γελάει συνέχεια χωρίς λόγο" This post is very toxic 4, because is hate language even if it does not contain offensive words
- 6.classify the given text based on toxicity levels. 1 is not toxic, 2 is slightly toxic, 3 is moderately toxic, 4 is very toxic, 5 is extremely toxic. text example: "Κόντεψα να κάνω εμετό από τα μάτια...αυτή γιατί γελάει συνέχεια χωρίς λόγο" This post is very toxic 4, because is hate language even if it does not contain offensive words. example text "pataw play.. variemai... variemai... variemai...telos video" which translates to "πατατω play.. βαριεμαι.. βαριεμαι.. βαριεμαι.. βαριεμαι.. τελος βιντεο, is a moderate toxic because it try to infer that the video is so bad that he is bored

**You**rate this text: "Ακούμε απο εσάς διαρκώς για ποιοτικές σχέσεις και σήμερα μας λέτε ότι το μοναδικό κριτήριο ενός ά<u>ντρα είναι η εμφάνιση!"</u>

## ⑥ ChatGPT

This text appears to express frustration or criticism about a perceived inconsistency in the messaging about quality relationships. Without additional context, the tone could be interpreted as slightly negative or critical. Therefore, I would classify it as moderately toxic. A rating of 3 (moderately toxic) seems appropriate based on the conveyed tone. Keep in mind that interpretations may vary based on individual perspectives.

- You you are rigth about the context and tone. Although would classify it as 2
- You rate this text: "Αβρραμμ φενεσαι πολύ σκληρός να τα τατουάζ.. " Ε Ε δεν ξέρω για σας εγω πάντως φοβήθηκα. Κάνε κανένα ακόμα στο κούτελο και έλα να μας το παίξεις σωστός και "υπερασπηστης" των αδικιμενων"

# ChatGPT

This text appears to contain playful banter and sarcasm. While it does involve some teasing, it doesn't seem to be explicitly toxic or harmful. The tone is more on the humorous side. Therefore, I would rate it as 1 (not toxic). Keep in mind that interpretations may vary, and understanding the context is crucial in assessing the perceived toxicity of a statement.

## 🕧 You

Αβρραμμ φενεσαι πολύ σκληρός να τα τατουάζ.. 🥮 😂 δεν ξέρω για σας εγω πάντως φοβήθηκα. Κάνε κανένα ακόμα στο κούτελο και έλα να μας το παίξεις σωστός και "υπερασπηστης" των αδικιμενων

this text contain high level or irony and criticizing. i would rate it as 3

- You

  rate this text" to video sou einai polu asxhmo"
- (§) ChatGPT

  The text "το video σου είναι πολύ ασχημο" translates to "your video is very ugly" in English. This statement is critical and negative, expressing a strong dislike for the video. Based on the negativity and the use of the term "ugly," I would classify it as moderately toxic. Therefore, a rating of 3 (moderately toxic) seems appropriate. Keep in mind that interpretations may vary based on context and individual perspectives.
- **You**i would rate it as 4 because he tries to bellitle the author, without providing context
  - You rate me this text: Mike μου αρέσουν σε ικανοποιητικό βαθμό τα επεισόδια που φτιάχνεις και τα παρακολουθώ ευχάριστα. Το μόνο που με στεναχωρεί είναι ότι δεν δίνεις αυτό το advantage και το motivational factor στο κοινό ώστε να πιστεύει ότι έχεις κάνει μια καλή έρευνα, έχεις διασταυρώσει πηγές και έχεις αφιερώσει χρόνο για την κατανόησή τους. Χωρίς να θέλω να υποτιμήσω τον κόπο που καταβάλλεις ώστε να προβάλεις και να παρουσιάζεις τέτοιου είδους videos.
    Φιλική συμβουλή: Προσπάθησε να βάζεις στην περιγραφή τις πηγές σου.
  - This text provides constructive feedback and criticism but is overall polite and offers suggestions for improvement. Given the constructive nature of the message, I would rate it as 1 (not toxic). It provides feedback on how the content creator can enhance their videos by including sources in the description, and it acknowledges the effort put into creating such content. It promotes a positive and helpful tone. If you have more
  - You
    that right

texts or questions, feel free to share!

### d You

rate me these texts:1. Κατάλαβα δε κατάλαβα , η μπυρα κατέβηκε

2.000 ναι Claude Shannon και Alan Turing τα θεμέλια της πληροφορικής

3.Δεν έχεις ακούσει ποτέ τον Σάνον

4.Μαλακα Μάικ έχει σκεφτεί ποτέ πόσα χρόνια έχουν περάσει από το πρωτο μπραφ?

5.Χάχα μας το είπε ο μαθηματικός στην Α λυκείου. Το θυμάμαι

6.Mikeius κανε ένα βιντεο για τον υπολογιστή των Αντικυθήρων

## 6 ChatGPT

Let's evaluate each text:

#### 1. "Κατάλαβα δε κατάλαβα , η μπυρα κατέβηκε"

 This text seems to express a light-hearted sentiment about understanding (or not understanding), possibly in a humorous context. It does not appear toxic. Rating: 1 (not toxic).

#### 2. "000 ναι Claude Shannon και Alan Turing τα θεμέλια της πληροφορικής"

 This text appears to express excitement or agreement about Claude Shannon and Alan Turing being foundational figures in computer science. It's positive. Rating: 1 (not toxic).

#### 3. "Δεν έχεις ακούσει ποτέ τον Σάνον"

 This text suggests someone hasn't heard of Shannon, possibly in a factual way. It doesn't seem toxic. Rating: 1 (not toxic).

### "Μαλακα Μάικ έχει σκεφτεί ποτέ πόσα χρόνια έχουν περάσει από το πρωτο μπραφ?"

This text uses informal language but doesn't appear explicitly toxic. It seems to
question whether Mike has thought about how many years have passed since the
first "braf" (possibly a term from the content). Rating: 1 (not toxic).

# 5. "Χάχα μας το είπε ο μαθηματικός στην Α λγκείου. Το θυμάμαι"

• This tend cominises a characteristic or an ablance in the control in high cabout later.

#### 5. "Χάχα μας το είπε ο μαθηματικός στην Α λυκείου. Το θυμάμαι"

 This text reminisces about something a math teacher said in high school. It's not toxic. Rating: 1 (not toxic).

#### 6. "Mikeius κανε ένα βιντεο για τον υπολογιστή των Αντικυθήρων"

 This text suggests making a video about the Antikythera mechanism. It appears informative and not toxic. Rating: 1 (not toxic).

Overall, these texts seem to have a positive or neutral tone and don't exhibit toxicity. If you have more texts or questions, feel free to share!

1 You

this text: "Δεν έχεις ακούσει ποτέ τον Σάνον" seems to belittling. i would rate it 2.

ChatGPT

ል ላ ጠ ን