# Sentiment Analysis on the comments of Pewdiepie videos

Rodrigo A. Chavez M. *i6165844* and Lillian A. Wu *i6151962*

May, 2019

**Abstract**

This report analyse the comments of 50 recent videos of Pewdiepie and we evaluate the polarity and toxicity leveraging libraries like TextBlob and the pre trained model BERT. As a result we find that comments towards Pewdiepie videos are generally positive and not as toxic as certain online articles would report.

## 1 Introduction

Pewdiepie has became the biggest English speaking channel on Youtube with currently more than 95 millions subscribers. With this huge amount of audience, it is interesting to see how the audience reacts to his videos. With sentiment analysis, it will be possible to indicate a polarity in the positivity of the comment section under the videos, and whether it is generally toxic or not. Since Pewdiepie has variant types of video where he covers different topics, such as Meme review, Pew news or monthly book club, the correlation between different topics and the positivity would also be explored.

## 2 Dataset

For this analysis we used two datasets, the first one was created by us and consists of 200,000 comments from Pewdiepie videos which were extracted using the public API of Youtube by requesting 4,000 comments per video of the 50 most recent ones at date 27/05/2019. This were comments directed to the video and not to other users in order to analyse the replies to the videos and not the interaction across users. In this dataset each video has a unique ID which can be used to access the video in Youtube by going to the link www.youtube.com/watch?v=ID. The second dataset was retrieved from Kaggle Jigsaw Unintended Bias in Toxicity Classification (JUBTC) competition, this data consist of sentences manually labeled with different kind of toxicities (”toxic”, ”severe_toxic”, ”obscene”, ”threat”, ”insult”, ”identity_hate”) and a target score between 0 and 1 which represent how toxic each sentences. Most of the data set consist of clean entries were no kind of toxicity was labeled and the rest is labeled with some kind of toxicity where every label seems to be a subset of toxic since most of the times it has some kind of toxicity is also labeled as toxic. Due to lack of computational power, in this analysis we just used the first 90,000 entries of the train set.

## 3 Experiment

Our procedure to analyse the comments sentiment polarity and toxicity was the former: For sentiment polarity we decided to compare the scores of the TextBlob and Pattern python libraries which use similar methodologies based on word frequencies and lookup table of individual polarity score per words. Even when this libraries don’t use state of the art algorithms, their simpler implementation allow fast sentiment classification and this results can be use as a baseline to compare the toxicity classification. The next part of the experiment consisted of fine tuning the Bidirectional Encoder Representations from Transformers (BERT-Base, Uncased) pre-trained model using pytorch with the JUBTC dataset from Kaggle on the GPU at Google Colab (K80, 1 hour, 12 GB VRAM). Once BERT was fine tuned we predicted the toxicity per comment and calculated the average toxicity per video.

## 4 Fine tuning BERT

In order to fine tune BERT we decided to use pytorch and the JUBTC dataset, our implementation avoids reinventing the wheel and is based on the kernel provided by the user ”yuval r” in the JUBTC competition at Kaggle. Since our implementation was bounded to Google Colab GPU constraints we used 80,000 entries for testing and 10,000 for validation instead of 1.8 million that the dataset is made of. The tokenizetion used the BertTokenizer from the pytorch_pretrained_bert library and we set the max sequence length to 128 in order to remain in the GPU ram boundaries. We then converted the values in target column to 1 if the value was higher than 0.5 or 0 if that was not the case. Then the Bert For Sequence Classification model from the bert-base-uncased file was loaded and created a BertAdam optimizer which is a closer implementation of the optimizer used by Google to train BERT. Finally we used binary cross entropy with logits as the loss function and trained the model in 1 epoch using batches of size 32 to keep the training in the time and memory boundaries. After multiple iterations since we were computational constraint we decided our accuracy score of 0.9086 in the validation set was good enough to proceed and analyse the comments of the Pewdiepie videos. As reference, the accuracy of our model would be positioned within the top 1,900 scores out of 2,300 in the JUBTC challenge where the first place has 0.94587 accuracy at date 27/05/2019.

## 5 Results

The results were produced by scoring each comment with the respective method and averaging this scores per video which is then are used to sort the videos.

Besides gameplays and travel vlogs, Pewdiepie videos usually can be categorized as follow,

- Meme review: where pewdiepie rates and explain memes.
- Cringe Tuesday: Pewdiepie reviews cringy videos of a same theme.
- Book review: Pewdiepie does monthly review on the books he read during the month.
- Pew news: discussion on one or multiple hot topics, could be a person or a current event.

### 5.1 Polarity

In Table 1, we listed the videos with the top 5 and worst 5 polarity among the 50 videos.

| Rank | Video ID | Category |
|---|---|---|
| 1 | qPnTTA8BC8A | book review |
| 2 | C2fRC55rA8w | travel vlog |
| 3 | PGbAWTqUuxQ | gameplay |
| 4 | QNLARCvIATo | travel vlog |
| 5 | OEUsKLW1th4 | gameplay |
| 46 | WOSC6uGtBFw | meme review |
| 47 | rdaQsl9jqmw | gameplay |
| 48 | wFxCAWqvmBE | meme review |
| 49 | zYZ1Fd7iH90 | cringe Tuesday |
| 50 | DCkydkdhL8M | meme review |

**Table 1:** Categorized videos with the top 5 highest and least polarity

Using both Pattern and TextBlob, the polarity values varied, but the top 5 and worst 5 ranking remains the same. TextBlob generally has a slightly higher polarity than Pattern.

| Rank | Pattern Polarity | TextBlob Polarity |
|---|---|---|
| 1 | 0.4933 | 0.4956 |
| 2 | 0.3267 | 0.3277 |
| 3 | 0.3185 | 0.3218 |
| 4 | 0.2999 | 0.3009 |
| 5 | 0.2640 | 0.2656 |
| 46 | 0.0935 | 0.0964 |
| 47 | 0.0901 | 0.0899 |
| 48 | 0.0628 | 0.0635 |
| 49 | 0.0581 | 0.0587 |
| 50 | 0.0422 | 0.0448 |

**Table 2:** Polarity comparison of the two libraries

### 5.2 Toxicity

In Table 3, we listed the videos with the top 5 and lowest 5 toxicity among the 50 videos. The list of the 50 videos sorted by toxicity can be found in Table 7 at the Appendices section.

## 6 Analysis

### 6.1 Polarity

There is clearly a contrast between the higher and lower polarity values. Book review videos has an overwhelmingly positive result. From the full results, we can conclude that travel vlogs tend to have higher polarity and meme reviews tend to have lower polarity. The polarity of a gameplay can differ drastically based on the game.

| Rank | Video ID | Categ. | Toxic |
|---|---|---|---|
| 1 | JLREgYXXdB8 | cringe Tue. | 0.2964 |
| 2 | eHYkTUmsJlY | Pew news | 0.1592 |
| 3 | JxAUHg8AguA | cringe Tue. | 0.1536 |
| 4 | 4QnLRnKwFM0 | Pew news | 0.1501 |
| 5 | 3m4mF9-7L-Y | Pew news | 0.1368 |
| 46 | rc1VR54nHV0 | collab. | 0.0612 |
| 47 | OEUsKLW1th4 | gameplay | 0.0604 |
| 48 | wFxCAWqvmBE | meme re. | 0.0522 |
| 49 | C2fRC55rA8w | travel vlog | 0.0498 |
| 50 | qPnTTA8BC8A | book re. | 0.0482 |

**Table 3:** Categorized top 5 toxic videos and least 5 toxic videos

### 6.2 Biased in the comments

There is a possibility that the result is not completely accurate since the in the comments that could be biased and potentially affect the polarity. For example, from Table 3 we compare two videos that are ranked 3 and 49 accordingly. The first video is the gameplay of ”Happy Wheel” and the second video is the gameplay of ”The Walking Dead”. The word ”happy” occurred a lot more times since it's part of the game name. And the word ”happy” adds to the polarity. On the contrary, the word ”dead” is in the name of the second game, thus there are plenty of comments using the word and could negatively affect its of polarity.

| Rank | ”happy” occurrence | ”dead” occurrence | Polarity (Pattern) |
|---|---|---|---|
| 3 | 438 | 130 | 0.3185 |
| 49 | 63 | 173 | 0.0581 |

**Table 4:** Specific word counts between two videos among its 4000 comments

### 6.3 Toxicity

Looking at Table 3 we can see how the model has classified as top 5 toxic videos belonging to the cringe Tuesday and Pew news categories, where the most toxic video ”I broke my ass” has almost double the toxicity score as the next video in the table. After deeper inspection on the comments in this video, we found out that the toxicity model was biased to score as toxic the comments that contained the word 'ass' which appears in 956 comments out of the 4000 (24% of the comments in that video). We found 'I Love you and your broken ass' as an example of a wrongly categorized comment where clearly the user intention was not toxic yet the model score it as a highly toxic with 0.9687. Furthermore the least toxic videos belong to a more diverse set of categories where the least toxic video 'Please watch for watch time thanks - Plato, The Republic' talks about philosophy books.

### 6.4 Correlation of Polarity and Toxicity

While toxicity and polarity are two different attributes we found that 5 of the top 10 positive (Table 5 in Appendices) videos are also in the top 10 least toxic videos (Table 6 in Appendices). Furthermore 4 of the most negative videos are in the top 10 most toxic videos. The difference in the top 10 list can be mainly explain due to the bias and the different focus of the algorithms where the polarity of a comment can be low if is sad while it could remains as not toxic.

## 7 Conclusion

Based on the results, we can conclude that generally the comments of Pewdiepie's videos are more positive than negative, and in 80% of the sample videos, less than 10% of the comments are toxic (Table 7). We also found out that the sentiment polarity and toxicity somewhat correlates in the top 10% percentile. Finally after analysing the results we discovered that the models weren't unbiased and further research is recommended.

## 8 Further research

We propose further research in tuning the BERT model to be less biased in order to get more reliable scores. We then suggest to compare the comments from Youtube to the comments from Reddit and see if the polarity and toxicity trend are similar or whether the platform type affects the way people express themselves. The results are not alarming and could be expected in other internet communities, we then propose to analyse other youtubers and see whether the critics of mainstream media are justifiable against Pewdiepie or if they are bias in some sort. Finally deeper statistical analysis taking the total number comments and views of the video into account could bring a more accurate insight in the behaviour of the viewers of Pewdiepie videos.

## Acknowledgments

## References

2018. Pewdiepie and the toxic culture of online gaming geek vs podcast.

Neha Bhangale. 2018. Toxic comment classification models comparison and selection.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.

Zhang W. Zhang M. Wu J. Wen T. Jiang, M. 2016. An LSTM-CNN attention approach for aspect-level sentiment classification. *Journal Of Computational Methods In Sciences And Engineering*, pages 1–10.

Narwal N. Khieu, K. 2018. *"Detecting and classifying toxic comments*. Ph.D. thesis, Iowa State University. https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/reports/6837517.pdf.

Amar Krishna. 2014. *"Polarity trend analysis of public sentiment on YouTube*. Graduate theses and dissertations, Iowa State University. https://lib.dr.iastate.edu/etd/13670.

F. Liu and X. Wu. 2018. Toxic comment detection with bi-directional lstm.

yuval r. 2014. Kernel for training bert using pytorch.

Aja Romano. 2018. Youtube's most popular user amplified anti-semitic rhetoric. again.

Alessandro Uryupina Olga Plank Barbara Filippova Katja Severyn, Aliaksei Moschitti. 2014. Opinion mining on youtube. https://www.researchgate.net/publication/263084636_Opinion_Mining_on_YouTube.

Jin Wang, Liang-Chih Yu, K. Robert Lai, and Xuejie Zhang. 2016. Dimensional sentiment analysis using a regional CNN-LSTM model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 225–230, Berlin, Germany. Association for Computational Linguistics.

## A Appendices

| Video ID | Polarity (Pattern) |
|---|---|
| DCkydkdhL8M | 0.0422 |
| zYZ1Fd7iH90 | 0.0581 |
| wFxCAWqvmBE | 0.0628 |
| rdaQsl9jqmw | 0.0901 |
| WOSC6uGtBFw | 0.0935 |
| MytxkcXV1Hk | 0.1005 |
| gvYwA_1T670 | 0.1099 |
| 96SZCm0RudE | 0.1103 |
| 4QnLRnKwFM0 | 0.1199 |
| JxAUHg8AguA | 0.1240 |
| rc1VR54nHV0 | 0.2234 |
| Ah5MYGQBYRo | 0.2299 |
| Pxgvgh9IFqA | 0.2313 |
| 3m4mF9-7L-Y | 0.2348 |
| vuuAMd2DaT4 | 0.2624 |
| OEUsKLW1th4 | 0.2640 |
| QNLARCvIATo | 0.2999 |
| PGbAWTqUuxQ | 0.3185 |
| C2fRC55rA8w | 0.3267 |
| qPnTTA8BC8A | 0.4933 |

**Table 5:** 10 lowest polarity and 10 highest polarity

| Rank | Video ID | Toxicity |
|---|---|---|
| 1 | JLREgYXXdB8 | 0.2964 |
| 2 | eHYkTUmsJlY | 0.1592 |
| 3 | JxAUHg8AguA | 0.1536 |
| 4 | 4QnLRnKwFM0 | 0.1501 |
| 5 | 3m4mF9-7L-Y | 0.1368 |
| 6 | DCkydkdhL8M | 0.1317 |
| 7 | PGbAWTqUuxQ | 0.1191 |
| 8 | zYZ1Fd7iH90 | 0.1185 |
| 9 | 5u8VYTQg4q8 | 0.1117 |
| 10 | ySyoESEU5NU | 0.1112 |
| 41 | QNLARCvIATo | 0.066 |
| 42 | gvYwA_1T670 | 0.0655 |
| 43 | yFBpOzvL-XA | 0.0645 |
| 44 | F_BQ9GO7Y88 | 0.0638 |
| 45 | w8ZRYdR6UuQ | 0.0623 |
| 46 | rc1VR54nHV0 | 0.0612 |
| 47 | OEUsKLW1th4 | 0.0604 |
| 48 | wFxCAWqvmBE | 0.0522 |
| 49 | C2fRC55rA8w | 0.0498 |
| 50 | qPnTTA8BC8A | 0.0482 |

**Table 6:** Top 10 toxic videos and least 10 toxic videos

| Rank | Video ID | Toxicity |
|---|---|---|
| 1 | JLREgYXXdB8 | 0.2964 |
| 2 | eHYkTUmsJlY | 0.1592 |
| 3 | JxAUHg8AguA | 0.1536 |
| 4 | 4QnLRnKwFM0 | 0.1501 |
| 5 | 3m4mF9-7L-Y | 0.1368 |
| 6 | DCkydkdhL8M | 0.1317 |
| 7 | PGbAWTqUuxQ | 0.1191 |
| 8 | zYZ1Fd7iH90 | 0.1185 |
| 9 | 5u8VYTQg4q8 | 0.1117 |
| 10 | ySyoESEU5NU | 0.1112 |
| 11 | Pxgvgh9IFqA | 0.1097 |
| 12 | Ux-YVt9iJLE | 0.1038 |
| 13 | rdaQsl9jqmw | 0.102 |
| 14 | E9vF215E5lw | 0.1 |
| 15 | bRG6sy3VaWU | 0.0973 |
| 16 | Ah5MYGQBYRo | 0.0968 |
| 17 | tN9vBO_b6b0 | 0.0956 |
| 18 | Li6c8n69bsU | 0.0886 |
| 19 | fsbZZT6dZIA | 0.0885 |
| 20 | opX0jijKojI | 0.0866 |
| 21 | gNNSlfcOBAY | 0.0856 |
| 22 | WOSC6uGtBFw | 0.0851 |
| 23 | dF7np7liv3o | 0.0843 |
| 24 | skFOyjFJ4pc | 0.0822 |
| 25 | 96SZCm0RudE | 0.082 |
| 26 | SFnMkigtHyI | 0.0794 |
| 27 | o3tSAKerrnU | 0.0787 |
| 28 | zc6iMpaYIUM | 0.0771 |
| 29 | 8TnOc5K9Spk | 0.0766 |
| 30 | LnO6Yqxx_M8 | 0.0764 |
| 31 | pNCkp6x5yT8 | 0.0754 |
| 32 | R6wzEuqdSdM | 0.0739 |
| 33 | xVVYkyZzCFY | 0.0706 |
| 34 | BATg388DiDQ | 0.0703 |
| 35 | cA-AT80GjxA | 0.0701 |
| 36 | vuuAMd2DaT4 | 0.0677 |
| 37 | wLcWf8fN54o | 0.0675 |
| 38 | MytxkcXV1Hk | 0.0675 |
| 39 | rNbNEcKpFz4 | 0.067 |
| 40 | QNLARCvIATo | 0.066 |
| 41 | gvYwA_1T670 | 0.0655 |
| 42 | yFBpOzvL-XA | 0.0645 |
| 43 | F_BQ9GO7Y88 | 0.0638 |
| 44 | w8ZRYdR6UuQ | 0.0623 |
| 45 | rc1VR54nHV0 | 0.0612 |
| 46 | OEUsKLW1th4 | 0.0604 |
| 47 | wFxCAWqvmBE | 0.0522 |
| 48 | C2fRC55rA8w | 0.0498 |
| 49 | qPnTTA8BC8A | 0.0482 |
| 50 | wKutrYtv3H0 | 0.047 |

**Table 7:** The 50 videos sorted by decreasing toxicity