Oversharing on Reddit: A Case Study of the Subreddit "AmItheAsshole"

Benites
Indiana University
Bloomington, United States
mbenites@iu.edu

***Abstract*—This document is a case study of the subreddit "AmItheAsshole", where we will identify the main topics discussed in the subreddit and analayze the community's reaction to these topics.**

***Index Terms*—NLP, social media, clustering, oversharing**

I. Introduction

A. The Relevance of the Case Study

As established by Leavitt [12] little research has examined the "role some social media platforms play in affording users temporary identities". The rise of social media has seen the emergence of anonymous accounts, accounts people don't want associated with themselves [12].

This phenomenon is clearly seen with the rise of the term 'throwaway accounts' in the platform of Reddit where people actively include the term in their username. Leavitt's research asserts the use of these 'throwaway' accounts is mainly used when the user discloses personal information or wishes to separate to 'spheres of his life'(work, friends, etc) [12].

Created in 2013 the subreddit1 'AmItheAsshole' is one of the most populars subreddits in Reddit with 22 million members. People in the subreddit generally post under 'throwaway accounts', seeking advice or the judgment of their actions detailed in their post. The posts disclose sensitive information like gender, age, personal thoughts and sensitive situations. With the use of anonymity the threat and discomfort caused by the posting of private or sensitive content [15] becomes invalid in the context of this subreddit.

According to Pew Research [13] center 18 to 29 year-olds made up 44 percent of Reddit users in 2024. In Brammer's study [3] oversharing is defined as sharing "too much information", the study sustained that topics often subjected to oversharing were medical issues, politics and relationships.

Examining the posts under this subreddit could allow us to study the phenomenon of oversharing in the 18 to 29 year olds age range as well as more a more detailed insight into what topics people feel the need to use an anonymous account, which becomes more interesting considering according to Pew Research center this age group also make up the majority users for Instagram, LinkedIn , X, Pinterest, Snapchat, Youtube, and TikTok.

B. Related Works

Graph based ranking is one of the many automatic keyphrase extraction techniques, Bougouin and Boudin[16] introduced TopicRank, with the goal of improving on Textrank. By making the topics the vertices they focus more on filtering and then clustering similar words to define a topic. Initially, clustering the top ranked words to automate the defining of topics was attempted; however, the results were not very accurate. Nevertheless the data used would have benefitted from further preprocessing and a larger training data.

Research by Basher and Sabri [15] focuses on generating comments that respond appropriately to the complex social and ethical considerations in the situation. The study received feedback on the generated comments by humans on. On average comments were correctly identified as human.

The model used for the study was 'Bart' introduced by Facebook, available to the public which brings into question the presence of artificially generated comments in the subreddit. Even if a thorough training process was used in the study, much simpler models might be used to generate content in the comments which may affect our measurement of the overall reaction to the post. For this reason we will extract the top 15 comments since they are less likely to have been artificially generated.

II. Methods

Web scraping techniques were used to extract Reddit posts from the last two weeks. A total of 849 posts were extracted using a Python script using Selenium Webdriver to scrape the subreddit. The post's url, score—-number of upvotes—-, author, post title, post id, post text content, and number of comments were collected from the subreddit's feed.

For each unique post encountered, the script followed the url to the post's comment page where the top comments were extracted once they loaded. For each comment the author, score, post id, comment id and comment content were extracted.

The data was stored in a SQLite database, containing 850 tables. 'AmItheAsshole' was the table containing the data collected from the feed page and the remaining tables, containing the comments for each post, were each named after their respective post's id and contained a foreign key which linked them to the 'AmItheAsshole' table.

To preprocess the text, the contractions were expanded to their full forms, the punctuation marks were removed, and the stop words from the NLTK stopwords list were removed. For each post, the present digits and their surrounding context were

examined. If they fell in the range of 13 to 90, their context was checked to determine whether it referenced age or if it contained the common expression 'I (age) gender'.

The preprocessed text was then converted into tensors, after the digits were removed, and converted into training data for the Word2Vec model which returned a vector of 70 dimensions for each word.

Hulth [17] states that most human-assigned keyphrases are nouns; therefore, using the NLTK 'pos tagger' the nouns per sentence within each post were extracted. sknetworks' PageRank was used to calculate the score for each word. From there, the seven highest ranked keywords were mapped into a new vector space given by a pre-trained Glove embeddings, 100 dimensions per word.

The new vector space was used since the specific context used for each word on Reddit and the semantic characteristics of the text did not prove useful for extracting a topic from the keywords. For example, the word 'classroom' was closest to 'relationship'.

Based on the most co–ocurrent keywords the topics were selected : 'family', 'boundaries', 'friends', 'money', 'relationship', 'health', 'work', 'stranger', 'living', 'pet' . To improve the quality of the results additional related words were assigned to the keywords, such as plural or more descriptive words such as 'boyfriend' or 'girlfriend' to obtain better results however we will be referring only to the words listed above as the topics.

The keywords, mapped in the new vector space, were assigned the closest topic vector using cosine similarity. Each score had weights, assigned based on the pagerank score of the keyword—-the higher the score the higher the weight—-and then the highest scoring word was selected as the topic.

For the comments, the automoderator entries were removed and the terms 'NTA' and 'YTA', standing for 'Not The Asshole' and 'You are The Asshole' respectively, were searched and annotated if present for each comment. Then applying VADER, a rule-based model built for classifying social media posts as negative, positive or objective, a score for each of the former assigned.

III. Results

As shown in Figure 1, the number of men and women using the platform is similar. Differences arise in age, specifically in the third quartile where data becomes more sparse for men up, extending to almost 50 years old. However, even this small random sample, the statistic proposed by Pew Research Center are supported.

The average sum of the selected keyword's PageRank scores was 0.35/1. As shown in figure 2 many of the keywords repeated across posts and across different positions. Many keywords also shared similar meanings.

As shown in figure 3 most of the posts were given 'family' as a topic. After mapping the keywords to the new space, one hundred posts's classification was manually verified and 64 out of 100 times the classification was accurate. The confidence interval ranging between 54.5 percent and 73 percent.

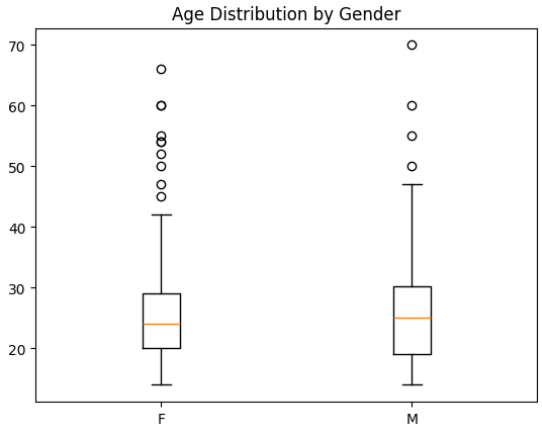

Fig. 1. Fig 1. Age distribution by gender for 64% of users who declared their age.

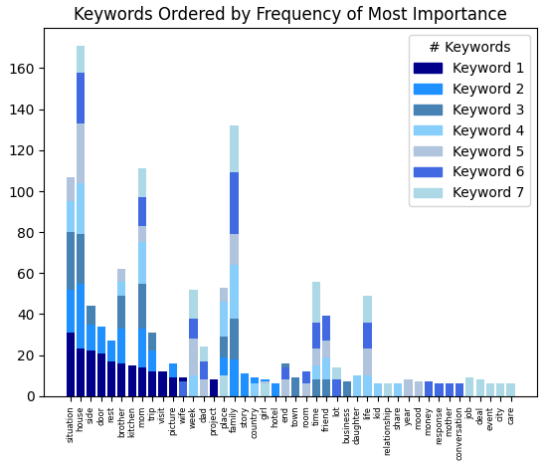

Fig. 2. Fig 2. Keyword frequency ordered by Frequency of highest scored keywords.

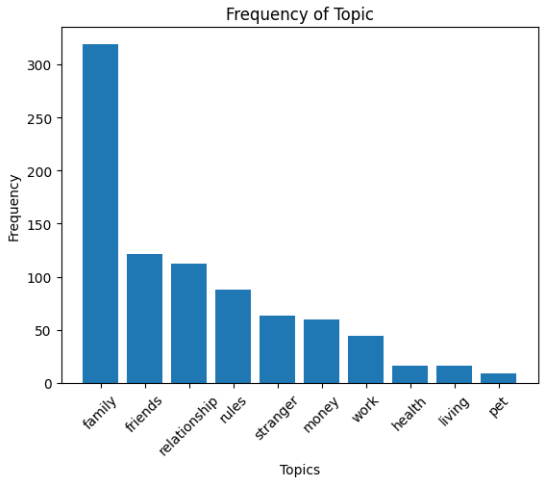

Fig. 3. Fig 3. Distribution of selected topics Across 849 entries.

When using VADER for classifying the sentiment of the comments, most were classified as neutral. However, perception could still be collected from the comments due to the nature of the subreddit where users classify the poster as 'an asshole' or 'not the asshole'. The most frequently mentioned per each comment(the mode) serves as the label for this attribute.

TABLE I
TOPIC-WISE CLASSIFICATION OF COMMENTS

Topic	**Attributes**		
	'Asshole'	***'Not the Asshole'***	***Age Range***
Family	54	256	66
Friends	20	96	50
Relationship	15	92	51
Rules	12	72	35
Stranger	11	50	46
Money	12	46	52
Work	9	34	26
Health	3	11	23
Living	5	12	30
Pet	3	6	33

IV. DISCUSSION

Based on the data, we can conclude that most of the posts are family-related, and that the majority of commenters adopt a neutral tone, using descriptive words rather than those VADER associated with emotions. Then again, this could also be because there was an unintentional oversampling of posts classified as 'Not the Asshole'. There was also limited data for the topics with less than 120 entries, with 849 posts collected the data size remains small, which is why the study would have benefited from further information extraction. Additionally, the data was very sparse in some parts and very dense in others, this is mostly due to the nature of social media, but it still makes it hard when you have a median of 500 and many outliers ranging from 10 000 to 20 000 which happened with the scores.

Using more advanced techniques for meaning and semantic extraction would also be beneficial since the lower bound of the confidence interval dropped down to 54.5 percent. Pretrained models such as RoBERTa or other transformer models would improve the study's result.

REFERENCES

[1] Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv. https://arxiv.org/abs/1301.3781

[2] Krbec, P. (2006). Language Modeling for Speech Recognition of Czech. https://www.semanticscholar.org/paper/Language-Modeling-for-Speech-Recognition-of-Czech-Krbec/4b40bed2f3f59e7476ba267cf4562cedbf210d39cited-papers

[3] Brammer, S. E., Punyanunt-Carter, N. M., Duffee, R. S. (2022). Oversharing on social networking sites: A contemporary communication phenomenon. Computers in Human Behavior Reports, 8, 100222. Elsevier. https://www.sciencedirect.com/science/article/pii/S2451958822000707

[4] Joshi, P. (2024, October 15). An Introduction to Text Summarization using the TextRank Algorithm. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2018/11/introduction-text-summarization-textrank-python/h-implementation-of-the-textrank-algorithm

[5] Zhu, L., Huang, M., Chen, M., Wang, W. (2021, September 7). An N-gram based approach to auto-extracting topics from research articles. arXiv. https://arxiv.org/abs/2110.11879

[6] Hutto, C., Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), 216-225. https://doi.org/10.1609/icwsm.v8i1.14550

[7] Yujun Wen, Hui Yuan and Pengzhou Zhang, "Research on keyword extraction based on Word2Vec weighted TextRank," 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, 2016, pp. 2109-2113,

[8] Zhu, L., Huang, M., Chen, M., Wang, W. (2021, September 7). An N-gram based approach to auto-extracting topics from research articles. arXiv. https://arxiv.org/abs/2110.11879

[9] Baumgartner, J. M., Lazzarin, E., Seiler, A. (2017). Pushshift API. GitHub. https://github.com/pushshift/api

[10] T. Mikolov, W.T. Yih, G. Zweig. Linguistic Regularities in Continuous Space Word Represen- tations. NAACL HLT 2013.

[11] Aldous, K. K., An, J., Jansen, B. J. (2022). What really matters?: characterising and predicting user engagement of news postings using multiple platforms, sentiments and topics. Behaviour Information Technology, 42(5), 545–568.

[12] Leavitt, A. (2015). "This is a Throwaway Account": Temporary technical identities and perceptions of anonymity in a massive online community. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work Social Computing (pp. 317–327). Association for Computing Machinery.

[13] Pew Research Center. (2024, November 13). Social Media Fact Sheet. Pew Research Center: Fact Sheets on Tech Adoption Trends.

[14] Taylor, T. L. (2019). World making or world breaking?: a Black womanist perspective on social media crises in higher education. Communication Education, 68(3), 381–385.

[15] Bsher, O., Sabri, A. (2023, October 21). AITA: Generating moral judgments of the crowd with reasoning.

[16] Bougouin, A., Boudin, F., Daille, B. (2013). TopicRank: Graph-Based Topic Ranking for Keyphrase Extraction. Proceedings of the International Joint Conference on Natural Language Processing (IJCNLP), 543–551.

[17] Anette Hulth. 2003. Improved Automatic Keyword Extraction Given More Linguistic Knowledge. In Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, pages 216–223, Stroudsburg, PA, USA. Association for Computational Linguistics.