

Using BERT to Understand TikTok Users' ADHD Discussion

Kayla Pineda
Computer Science
Old Dominion University
Norfolk, United States
kpine003@odu.edu

Anne M. Perrotti
Communication Disorders and
Special Education
Old Dominion University
Norfolk, United States
APerrott@odu.edu

Faryaneh Poursardar
Computer Science
Old Dominion University
Norfolk, United States
poursardar@cs.odu.edu

Danielle Graber
Computer Science
Christopher Newport University
Newport News, United States
danielle.graber.19@cnu.edu

Sampath Jayarathna
Computer Science
Old Dominion University
Norfolk, United States
sampath@cs.odu.edu

Abstract—The rapid rise in popularity of the social media platform TikTok has allowed information to reach a wider audience at a faster rate compared to traditional news outlets. With this growth, more young people are connected with each other, allowing discussion on issues of mental health, a once stigmatized topic, to grow among young people. However, not all information about mental health issues is done properly or accurately. A commonly discussed mental health issue on TikTok is Attention-Deficit/Hyperactivity Disorder, known as ADHD. In this paper, we briefly describe how ADHD is portrayed on TikTok. Then, we analyze ADHD discussion around themes of self-diagnosis, symptoms, and self-help through text analysis using BERT. We analyze the common responses to ADHD TikToks and set up a multi-label classifier to understand the general range of responses.

Index Terms—ADHD, TikTok, web scraping, BERT, multi-label classification

I. INTRODUCTION

One major topic of discussion among young people is mental health. Mental health is becoming destigmatized thanks to the internet. People want to explore mental health, so they use different online platforms, especially social media, to discuss it. However, not all mental health discussion is representative, objective, or beneficial. Like any chain of communication, a topic that is shared between a vast group of people is bound to be translated, obscured, and changed to create disinformation.

A. TikTok's Rise in Popularity

In this exploratory study, we focused on TikTok due to its quick rise in popularity and audience. It was ranked the 7th most used social media app worldwide in 2021 [1]. The largest demographic at 25% are users between ages 10-19, and the second largest demographic being users ages 20-29 at 22.4% [1]. TikTok is a popular, well-established, and easily accessible platform, making it meaningful to investigate.

B. Mental Health Discussion on TikTok

As of July 7th, 2022, the mental health hashtag had 39 billion uses [2]. People are accessing TikTok for mental health advice, awareness, symptoms, connection, and advocacy. There are several mental health diagnoses trending on TikTok. In this paper, we limit our scope to ADHD. As of July 2022, there were 12.7 billion uses of the #adhd tag [3].

C. ADHD Portrayal on TikTok

The majority of ADHD related TikToks do not come from medical professionals. Instead, most TikToks convey users' personal experiences with ADHD by sharing symptoms, actions, or scenarios possibly caused by their ADHD.

ADHD is a complex mental health diagnosis manifesting itself differently across people [4]–[7], making it difficult for TikToks to capture or represent ADHD with validity and reliability. What may be true for the user who posted the TikTok, may not be true for everyone. Our main aim in this paper is to understand users' responses to ADHD TikToks communicating diagnostic information. We plan to use this in further research on the existence and effect of misinformation and disinformation about mental health diagnoses on social media platforms

II. METHODOLOGY

To understand the reaction to ADHD diagnostic TikToks, the data collection and analysis process was split into three segments: scraping of TikTok comments, labeling of the comments, and training a multi-label classifier using BERT.

A. Web Scraping

The first step was collecting comments through web scraping. We scraped 12 diagnostic TikToks between January 17, 2023 and February 9, 2023. These TikToks were randomly selected using the search term "#ADHD." After scraping the comments, we removed non-English and duplicate comments.

Then, we limited the comments to those with specific keywords: ADD, ADHD, Anxiety, Depression, Diagnosis, Self-diagnosis, Symptoms, and TikTok. These keywords were selected because they represent diagnostic themes communicated through the TikTok content. We ended with a total of 1,327 comments using the above criteria.

B. Labeling

Next, we labeled the comments using the keywords mentioned above.

1) *Labeling Criteria:* The first criteria applied indicated explicit use or a direct statement of the word in the comment, and the second criteria applied indicated implicit or implied meaning of the word as mentioned in the comment. For example, the comment "Maybe get tested because it's normal to do this but also people with ADHD do it too" would get the labels ADHD, Diagnosis, and Symptoms. The comment gets the label ADHD as they explicitly use the word. For Diagnosis, the user mentions getting tested, alluding to a diagnosis. For Symptoms, the user says "it's normal to do this" which references the symptoms mentioned in the TikTok.

2) *Labeling Process:* We had three reviewers. The first reviewer labeled all comments according to the defined, objective criteria. The second reviewer went through the labeled comments and commented if they disagreed with the label(s) given. The first reviewer then reviewed the comments again. If the first reviewer agreed with the change the second reviewer recommended, that comment was resolved. At the time of this writing, the third reviewer is in the process of reviewing the comments. For our analysis, we use a subset comprised of 940 comments that yielded 100% inter rater reliability.

C. Multi-label Classification

Using a multi-label classifier with BERT [8], our plan was to train a custom classifier to identify common responses to ADHD related TikToks. For our preliminary work, we did an 80% training 20% testing split with the 940 comment subset.

III. RESULTS

A. Labeling Counts

For the labels, we counted the number of instances for various combinations. In Table I we show the top five combinations.

TABLE I
TOP FIVE LABEL COMBINATIONS

Label Combination	Count	Percentage
ADHD, Self-diagnosis	176	18.72%
ADHD, Diagnosis	169	17.98%
Diagnosis, Symptoms	160	17.02%
ADHD, Symptoms	134	14.26%
Self-diagnosis, Symptoms	93	9.89%

B. Classifier Training

For our preliminary training, we received the following results:

TABLE II
CUSTOM CLASSIFIER TRAINING RESULTS

Accuracy (Average)	59.36%
ROC Accuracy (Average)	90.42%
F1 Score (Average)	86.20%

IV. DISCUSSION

The most common label combination was ADHD and Self-diagnosis, meaning the most common responses were users believing they have ADHD and opting to self-diagnose, or that users discourage self-diagnosing. The second common combination was ADHD and Diagnosis, meaning users agreed with the TikTok because they have an official ADHD diagnosis, or users desire to get an official diagnosis.

The remaining three combinations include the symptom label, indicating that users related to the symptoms mentioned or demonstrated in the TikTok post. Users linked these symptoms with a formal diagnosis, a self-diagnosis, or to the common general understanding of ADHD. Often, the representation of symptoms depicted basic outward behaviors often associated with ADHD "like" behavior versus objective measures of cognitive or physiologic brain function.

V. CONCLUSION

Mental health discussion has become popular on TikTok, especially regarding the mental health diagnosis ADHD. Many users suspect they may have ADHD based on a posted short-form video. However, not all short-form videos capture or represent objective, valid, and reliable diagnostic criteria for ADHD. ADHD is a layered mental health diagnosis, so it is important for TikTok users to do their research to further their understanding of ADHD and use professional guidance in seeking a diagnosis or treatment.

In the future, we hope to broaden the scope of this research to the mental health diagnosis Autism Spectrum Disorder (ASD). We want to quantify disinformation regarding ADHD and ASD on TikTok and understand how this disinformation impacts user perspectives and advocacy.

REFERENCES

- [1] B. Dean, "TikTok user statistics (2022)," 2022.
- [2] "#mentalhealth hashtag videos on tiktok," 2022.
- [3] "#adhd hashtag videos on tiktok," 2022.
- [4] L. C. Klein and E. J. Corwin, "Seeing the unexpected: how sex differences in stress responses may provide a new perspective on the manifestation of psychiatric disorders," *Current psychiatry reports*, vol. 4, no. 6, pp. 441–448, 2002.
- [5] J. Gershon and J. Gershon, "A meta-analytic review of gender differences in adhd," *Journal of attention disorders*, vol. 5, no. 3, pp. 143–154, 2002.
- [6] R. Gomez, V. Stavropoulos, A. Vance, and M. D. Griffiths, "Gifted children with adhd: how are they different from non-gifted children with adhd?," *International Journal of Mental Health and Addiction*, vol. 18, pp. 1467–1481, 2020.
- [7] S. S. Agha, S. Zammit, A. Thapar, and K. Langley, "Are parental adhd problems associated with a more severe clinical presentation and greater family adversity in children with adhd?," *European child & adolescent psychiatry*, vol. 22, pp. 369–377, 2013.
- [8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.