

Original Paper

# Exploring Eating Disorder Topics on Twitter: Machine Learning Approach

Sicheng Zhou<sup>1</sup>, MSc; Yunpeng Zhao<sup>2</sup>, MSc; Jiang Bian<sup>2</sup>, PhD; Ann F Haynos<sup>3</sup>, PhD; Rui Zhang<sup>1,4</sup>, PhD

<sup>1</sup>Institute for Health Informatics, University of Minnesota, Minneapolis, MN, United States

<sup>2</sup>Department of Health Outcomes & Biomedical Informatics, University of Florida, Gainesville, FL, United States

<sup>3</sup>Department of Psychiatry, University of Minnesota, Minneapolis, MN, United States

<sup>4</sup>Department of Pharmaceutical Care & Health Systems, University of Minnesota, Minneapolis, MN, United States

**Corresponding Author:**

Rui Zhang, PhD

Institute for Health Informatics

University of Minnesota

8-100 Phillips-Wangensteen Building

516 Delaware Street SE

Minneapolis, MN, 55455

United States

Phone: 1 612 626 4209

Email: [zhan1386@umn.edu](mailto:zhan1386@umn.edu)

## Abstract

**Background:** Eating disorders (EDs) are a group of mental illnesses that have an adverse effect on both mental and physical health. As social media platforms (eg, Twitter) have become an important data source for public health research, some studies have qualitatively explored the ways in which EDs are discussed on these platforms. Initial results suggest that such research offers a promising method for further understanding this group of diseases. Nevertheless, an efficient computational method is needed to further identify and analyze tweets relevant to EDs on a larger scale.

**Objective:** This study aims to develop and validate a machine learning–based classifier to identify tweets related to EDs and to explore factors (ie, topics) related to EDs using a topic modeling method.

**Methods:** We collected potential ED-relevant tweets using keywords from previous studies and annotated these tweets into different groups (ie, ED relevant vs irrelevant and then promotional information vs laypeople discussion). Several supervised machine learning methods, such as convolutional neural network (CNN), long short-term memory (LSTM), support vector machine, and naïve Bayes, were developed and evaluated using annotated data. We used the classifier with the best performance to identify ED-relevant tweets and applied a topic modeling method—Correlation Explanation (CorEx)—to analyze the content of the identified tweets. To validate these machine learning results, we also collected a cohort of ED-relevant tweets on the basis of manually curated rules.

**Results:** A total of 123,977 tweets were collected during the set period. We randomly annotated 2219 tweets for developing the machine learning classifiers. We developed a CNN-LSTM classifier to identify ED-relevant tweets published by laypeople in 2 steps: first relevant versus irrelevant ( $F_1$  score=0.89) and then promotional versus published by laypeople ( $F_1$  score=0.90). A total of 40,790 ED-relevant tweets were identified using the CNN-LSTM classifier. We also identified another set of tweets (ie, 17,632 ED-relevant and 83,557 ED-irrelevant tweets) posted by laypeople using manually specified rules. Using CorEx on all ED-relevant tweets, the topic model identified 162 topics. Overall, the coherence rate for topic modeling was 77.07% (1264/1640), indicating a high quality of the produced topics. The topics were further reviewed and analyzed by a domain expert.

**Conclusions:** A developed CNN-LSTM classifier could improve the efficiency of identifying ED-relevant tweets compared with the traditional manual-based method. The CorEx topic model was applied on the tweets identified by the machine learning–based classifier and the traditional manual approach separately. Highly overlapping topics were observed between the 2 cohorts of tweets. The produced topics were further reviewed by a domain expert. Some of the topics identified by the potential ED tweets may provide new avenues for understanding this serious set of disorders.

(*JMIR Med Inform* 2020;8(10):e18273) doi: [10.2196/18273](https://doi.org/10.2196/18273)

**KEYWORDS**

eating disorders; topic modeling; text classification; social media; public health

## Introduction

### Background of Eating Disorders and Social Media

Eating disorders (EDs) are a prevalent type of mental illness affecting more than 30 million people across different age groups in the United States [1]. These disorders are commonly underdiagnosed and undertreated [2], and even among individuals who receive diagnosis and treatment, recovery takes a long time to achieve and remains elusive to many [3,4]. Unfortunately, there are serious consequences associated with EDs. Affected individuals often experience significant negative psychological, physical, and interpersonal effects of ED symptoms [5]. Although evidence-based interventions are available for EDs, they are not helpful for many, suggesting that they may not be targeting the correct psychological variables for these individuals [6]. Thus, it is important to gather additional information on the thoughts, emotions, and behaviors of individuals with EDs to identify treatment targets to improve or develop effective interventions for these populations [7].

During the past decade, the number of users of social media platforms, such as Twitter and Facebook, has increased sharply. These platforms provide the general public with opportunities to express their thoughts and opinions and share information about their daily lives, including their health information. This practice has yielded a large number of social media messages that may provide valuable information on a variety of health topics. The analysis of these messages could produce knowledge, permitting more sensitive and accurate education and intervention design in different areas of public health [8]. As such, applying data mining techniques to analyze Twitter data has become a popular methodological approach in health care research.

For instance, a study in 2011 used the Ailment Topic Aspect Model, incorporated with previous knowledge, to create structured disease information from tweets that was subsequently used for the surveillance of a series of different ailments [8]. In 2016, Xu et al [9] checked the frequency of discussions on cancer-related topics among Twitter users and found differences among different race and ethnicity groups. In 2019, Musaev et al [10] applied a latent Dirichlet allocation (LDA) topic model to tweets to examine public discussions about cardiovascular disease and found that state health departments play an important role in communicating with the public about cardiovascular health. We have also conducted a series of studies using Twitter data on various health-related topics, from detecting adverse events to studying laypeople's discussion of human papillomavirus vaccination [11-17]. We have also conducted initial work on using Twitter data to identify discussion topics relevant to EDs. These studies indicate that the analysis of tweets on a particular health topic, coupled with information derived from user profiles, may facilitate novel knowledge discovery in these areas of medicine. EDs have become a popular topic on Twitter. Thus, data mining methods and tools that can more effectively and efficiently identify and analyze these tweets and

associated user profiles may help advance the research on ED-related content on social media. In addition, these methods can be translated for use in research on social media use relevant to health-related topics.

### ED-Related Research and Gaps

Due to the self-protective nature of EDs, many individuals may not be willing to communicate about their experience of the disorder with others, potentially limiting the ability of researchers and clinicians to understand the factors that promote ED symptoms [18,19]. However, many individuals with an ED use social media to engage in a more open discourse about ED content with others with shared experiences [20]. Although data on Twitter are publicly available, few studies have explored public discussions about EDs. One study investigated how some Twitter accounts promote ED symptoms and the associated negative health consequences among Twitter users. They manually collected data from 45 ED-promotion Twitter accounts, including the profile information, the tweets posted by these accounts, and information about their followers. Through content analysis, they identified a list of ED-related keywords in these tweets and found a positive correlation between the percentages of ED-relevant tweets posted by the ED-promotion accounts and their followers [19]. Another study collected and reviewed ED-relevant tweets and manually classified the collected tweets into different subgroups to provide insights on EDs and to inform future web-based interventions for EDs [21]. These studies indicate that analysis of ED-relevant tweets may help to provide insight into factors that motivate ED behaviors, which may further be used to prevent and treat EDs. However, these studies mainly used keyword-searching strategies to collect ED-relevant tweets and analyzed the content of tweets through manual review. This approach is limited because it only permits analysis of a relatively small number of tweets within a limited time frame with compromised efficiency in content analysis. As a result, the obtained information may not be sufficiently comprehensive, and manual analyses may not be scalable. To improve these studies, computational methods are needed to identify and analyze ED-relevant tweets.

### Objective of the Study

To expand upon previous research on social media engagement among individuals with EDs, the focus of this study is to develop an automatic approach to better understand public perceptions and thoughts about EDs and ED-related behaviors using Twitter data. Specifically, a machine learning approach was developed to identify ED-relevant tweets, and a topic modeling method was implemented to analyze the content of the identified tweets. Potential ED-related factors, such as behaviors, thoughts, and mental status, were summarized through content analysis.

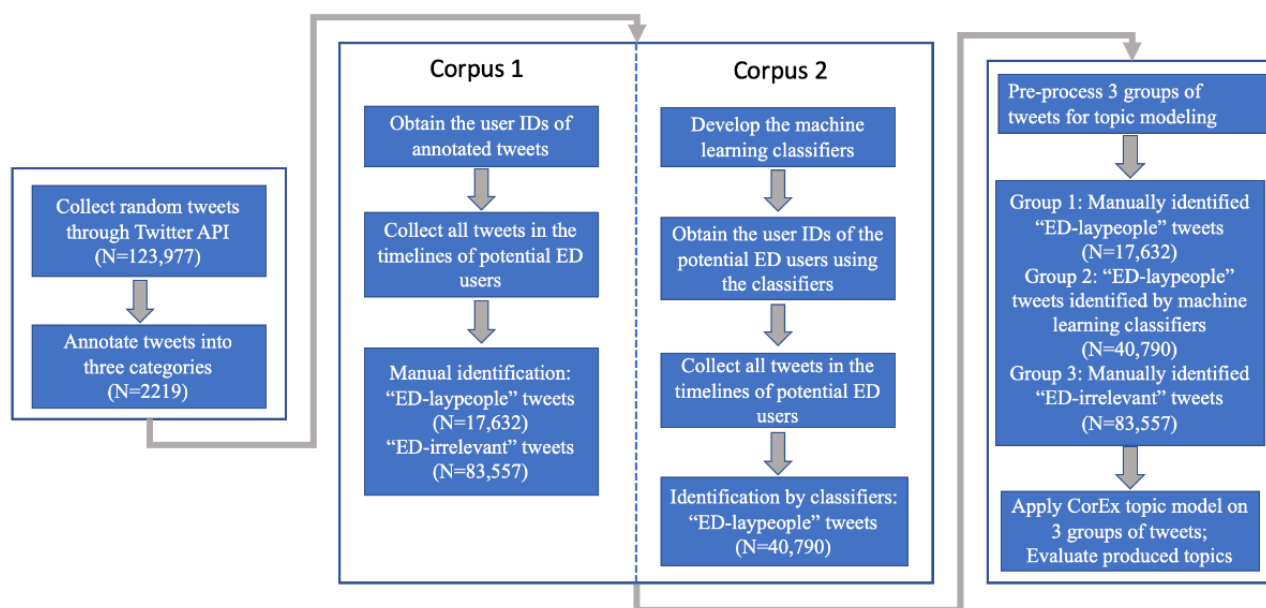
## Methods

### Overview of Experimental Pipeline

The overall experimental pipeline is shown in [Figure 1](#). We used 2 data sets spanning September 2012 to October 2019. We then randomly selected and annotated 2219 tweets to develop a training data set for the machine learning tasks of identifying ED-relevant tweets. We explored different machine learning-based methods to filter out ED-irrelevant tweets and classified the remaining tweets into either promotional and educational information or laypeople's discussions. We determined if a tweet was published by a *layperson* on the basis

of the content of the tweets. For example, if the tweet content was purely about an advertisement, the tweet was considered likely to be published by a *nonlayperson*. Thereafter, we performed topic modeling on 2 corpora. One corpus was built manually on the basis of 2219 annotated tweets and regarded as the gold standard for tweets that contain laypeople's discussions about EDs. Another corpus of ED-relevant tweets was built using the developed machine learning classifiers. We evaluated the topic modeling results to validate whether the corpus built by our developed classifiers could produce similar ED-relevant topics compared with the manually identified gold standard corpus. We also further identified and analyzed topics from the ED-irrelevant tweets posted by the potential ED users.

**Figure 1.** The workflow of the study. N represents the number of tweets at the corresponding step. CorEx: Correlation Explanation; ED: eating disorder.



### Twitter Data Collection

The Twitter data used in this study were from 2 different sources: (1) we used a list of ED keywords we developed from previous studies [19,21] ([Textbox 1](#)) and used these keywords to search for potential ED-relevant tweets from a database of random tweets collected from January 1, 2012, to September

30, 2018, using the Twitter streaming application programming interface (API); and (2) we used the same list of keywords to collect more Twitter data by using the Twitter search API from September 26, 2019, to October 30, 2019. This time span was selected because we have been collecting Twitter data since 2012.

**Textbox 1.** List of eating disorder–relevant keywords refined from previous studies.

<p>Eating disorder–relevant keywords:</p> <ul style="list-style-type: none"> <li>• Anorexia</li> <li>• Anorexic</li> <li>• Anamia</li> <li>• ana/mia</li> <li>• Anahelp</li> <li>• Anabuddy</li> <li>• Anaprobs</li> <li>• Binge</li> <li>• Bulimia</li> <li>• bulimic</li> <li>• #BingeEatingDisorder</li> <li>• #CompulsiveEating</li> <li>• ednos</li> <li>• edlogic</li> <li>• edprobs</li> <li>• edproblems</li> <li>• pro ana</li> <li>• proana</li> <li>• pro mia</li> <li>• promia</li> <li>• purge</li> </ul>
--

## Identification of Target Tweets

### Annotation of Tweets

Of the 123,977 total tweets, we randomly selected 2219 tweets on the basis of keyword distribution for gold-standard data set development. The tweets were annotated into 3 categories: *ED-irrelevant*, *ED-promotional and education*, and *ED-laypeople*. *ED-irrelevant* tweets were tweets considered irrelevant to EDs, *ED-promotional and education* tweets were tweets considered to be published by companies and institutes to promote their products or educate the public about EDs, and *ED-laypeople* tweets were considered to be ED-relevant tweets posted by individual users. The tweets labeled with *ED-laypeople* were our target tweets, and the individual users who posted the *ED-laypeople* tweets were defined as potential ED users. Owing to the nature of social media, it is extremely difficult to determine which user does indeed have an ED. However, a large portion of these users were highly engaged in the discussion of ED symptoms or frequently posted their activities and thoughts about EDs and how ED symptoms affected their lives. Two annotators started with 100 tweets to develop an initial annotation guideline to identify the category of each tweet on the basis of its content. Thereafter, we annotated another 100 tweets to refine this guideline. Finally, each of the remaining 2019 tweets were annotated by 2 coders.

Agreements were calculated, and conflicts were resolved through group discussions.

### Manually Identified ED-Laypeople Tweets and ED-Irrelevant Tweets

Within the 2219 tweets, we extracted the Twitter user accounts that posted tweets in the *ED-laypeople* category. We manually reviewed the usernames and parts of their tweets and removed users whose usernames and tweets indicated that they were not *laypeople*; for example, if a username contained the name of a company or an institute, the tweets from that account would be classified as not belonging to the *ED-laypeople* category. Through this process, 31 accounts were removed for being companies or institutes. The remaining accounts were classified as potential ED users. We then collected all the tweets posted by these potential ED users to construct their Twitter timelines. We checked all of their tweets to see if these tweets contained one of the ED keywords in [Textbox 1](#). If so, the tweets were regarded as *ED-laypeople* tweets, otherwise as *ED-irrelevant* tweets. The *ED-laypeople* tweets and the *ED-irrelevant* tweets were further analyzed using a topic modeling method. Although these *ED-irrelevant* tweets were classified as not directly associated with EDs, they were considered to reflect general topics within the potential ED users' daily lives, which could help to differentiate ED-related experiences from other aspects of these users' lives.

### Machine Learning Classifier Development

Supervised text classifiers can learn patterns from annotated input samples and automatically classify tweets into desired categories. To develop the classifiers, we first preprocessed the annotated tweets by replacing (1) hyperlinks (eg, *http://t.co/xxxx*) with *<url>*, (2) mentions (eg, *@username*) with *<user>*, (3) hashtags (eg, *#eatdisorder*) with *<hashtag> eatdisorder*, and (4) emojis with *<emojies>*. Then we explored 2 supervised deep learning algorithms (ie, convolutional neural network [CNN] and long short-term memory [LSTM]) and 5 machine learning algorithms (ie, naïve Bayes [NB], linear regression, support vector machine [SVM], random forest [RF], and gradient boosting trees [GB]) to classify our large tweet corpus automatically. We developed our classifiers in a 2-step process using the annotated tweets so that each classifier produced a binary output. In the first task, classifiers were trained to distinguish between ED-relevant and ED-irrelevant tweets (*ED-irrelevant* vs the union of *ED-promotional and education* and *ED-laypeople*) to filter out irrelevant tweets. In the second task, classifiers were trained to distinguish between the *ED-promotional and education* and *ED-laypeople* tweets. We developed a CNN classifier for the first task. The architecture of the CNN model included an embedding layer, a convolutional layer, a global max pooling layer, and a sigmoid output dense layer. We initialized the embedding layer with the Global Vectors for Word Representation pretrained 200-dimension Twitter word embeddings. In the convolutional layer, we set the number of filters to 64, the length of filters to 3, and the dropout rate to 0.2. For the second task, we developed an LSTM model. The architecture of the LSTM model included an embedding layer, an LSTM layer, a global max pooling layer, and a sigmoid output dense layer.

### Identifying Topics From Tweets of Potential ED Users

To understand the mental status and everyday life of potential ED users, we applied topic modeling to explore their tweets. Topic modeling has been a popular method for identifying latent patterns of words in a large collection of documents [22]. The most representative method for topic modeling is LDA—a probabilistic generative model [23]. In LDA, each document is assumed to contain a mixture of topics, where each topic is a probability distribution over the words in the document [24]. Some new topic models were developed to solve some of the limitations in LDA, such as the Biterm Topic Model (BTM) and the Correlation Explanation (CorEx) model [24,25]. BTM mainly improves the LDA's problem of sparse word co-occurrence patterns at the document level; thus, it uses the term *co-occurrence patterns* in the entire corpus to learn topics [24]. The CorEx model does not have an assumption about how the underlying data are generated, similar to LDA, which avoids assigning the characteristics of topics ahead of time. The CorEx model identifies the topics that are *maximally informative* about a collection of documents [25]. The BTM and the CorEx model were tested in our preliminary study [17], and the CorEx model was adopted because it produced more meaningful topics using our collected tweets.

We implemented the CorEx model on 3 groups of Twitter data: (1) the *ED-laypeople* tweets identified through manually curated rules posted by potential ED users, (2) the *ED-irrelevant* posted by potential ED users, and (3) *ED-laypeople* tweets identified by the machine learning algorithm, as mentioned earlier. For each group, we tested the CorEx model with different numbers of topics (n=50, 60, 70, 80, 90, 100). Although quantitative metrics are used to infer a reasonable number of topics (eg, perplexity and coherence), they sometimes cannot identify the optimal number of topics. On the basis of our experience in a previous study [17], we manually reviewed the topics produced by the different experiments to determine the optimal number of topics for further topic evaluation.

### Topic Evaluation

The results of the topic modeling experiments were further reviewed and analyzed by the domain expert (AH). Three steps were taken to evaluate the topic modeling results. First, the domain expert summarized the theme for each topic on the basis of the topic keywords. Second, on the basis of the top 10 most-relevant tweets for each topic, the expert judged whether each tweet was coherent with the summarized topic theme. The coherence rate of each topic, defined as the percentage of coherent tweets per topic, was calculated. Finally, topic themes with similar meanings were merged into higher-level categories.

## Results

### Tweet Collection and Annotation

Two coders annotated 2219 tweets into 3 classes, as mentioned earlier. Within 2219 tweets, 669 tweets were annotated as *ED-irrelevant*, 579 tweets were annotated as *ED-promotional and education*, and 971 tweets were annotated as *ED-laypeople*. The interrater agreement score between the 2 annotators was 0.84 on the basis of the first 200 tweets. We used the Cohen kappa test to calculate the score.

### Identification of the Target Tweets

#### Manual Identification of ED-Laypeople and ED-Irrelevant Tweets

As described earlier, we manually identified 17,632 *ED-laypeople* tweets and 89,312 *ED-irrelevant* tweets posted by the potential ED users.

#### Machine Learning Classifier Development

As mentioned earlier, 7 classifiers were explored (ie, CNN, LSTM, NB, LN, SVM, RF, and GB). We developed our classifiers in a 2-step process (ie, *ED-irrelevant* vs the other 2 labels and then *ED-promotional and education* vs *ED-laypeople*). Overall, 79.99% (1775/2219) of the tweets were used as the training set, and 20.01% (444/2219) tweets were used for evaluation. Table 1 shows the performances of the classifiers.

**Table 1.** Performances of the developed classifiers.

Classifier	Precision	Recall	F <sub>1</sub> score	P value
<b><i>ED<sup>a</sup>-irrelevant versus other 2 labels<sup>b</sup></i></b>				
CNN <sup>c</sup>	0.88	0.89	0.89	N/A <sup>d</sup>
LSTM <sup>e</sup>	0.86	0.89	0.88	.15
NB <sup>f</sup>	0.85	0.73	0.75	<.001
LN <sup>g</sup>	0.84	0.78	0.81	<.001
SVM <sup>h</sup>	0.87	0.83	0.85	<.001
RF <sup>i</sup>	0.86	0.85	0.86	.005
GB <sup>j</sup>	0.77	0.75	0.76	<.001
<b><i>ED-promotional and education versus ED-laypeople<sup>k</sup></i></b>				
LSTM	0.90	0.89	0.90	N/A
CNN	0.87	0.87	0.87	.006
NB	0.80	0.74	0.76	<.001
LN	0.83	0.80	0.81	<.001
SVM	0.82	0.79	0.80	<.001
RF	0.84	0.82	0.83	<.001
GB	0.84	0.82	0.83	<.001

<sup>a</sup>ED: eating disorder.

<sup>b</sup>ED-irrelevant versus other 2 labels: in this task, the performances of CNN and LSTM have no significant difference; they are both significantly higher than the others ( $P<.01$ ).

<sup>c</sup>CNN: convolutional neural network.

<sup>d</sup>N/A: not applicable.

<sup>e</sup>LSTM: long short-term memory.

<sup>f</sup>NB: naïve Bayes.

<sup>g</sup>LN: linear regression.

<sup>h</sup>SVM: support vector machine.

<sup>i</sup>RF: random forest.

<sup>j</sup>GB: gradient boosting trees.

<sup>k</sup>ED-promotional and education versus ED-laypeople: in this task, the performance LSTM is significantly higher than the others ( $P<.01$ ).

In the first task, the CNN outperformed the other classifiers ( $F_1$  score=0.89). The CNN classifier identified 88,261 tweets that were ED-relevant. In the second task, LSTM obtained the best performance ( $F_1$  score=0.90). Thus, we adopted LSTM for the second task. The LSTM method identified 40,790 *ED-laypeople* tweets posted by 21,600 Twitter users.

### CorEx Topic Model Implementation

The CorEx topic model was implemented on 3 groups of Twitter data, as mentioned earlier. After preprocessing, the first group

(*ED-laypeople* tweets manually identified) contained 17,632 tweets. The second group (*ED-laypeople* tweets identified by the developed CNN-LSTM classifier) contained 40,790 tweets. The third group (*ED-irrelevant* tweets posted by the potential ED users) contained 83,557 tweets. After the initial review, the optimal number of topics was determined to be 70 for groups 1 and 2. For group 3, the optimal number of topics was determined to be 80. [Textbox 2](#) shows the representative words of selected topics obtained from 3 groups of tweets.

**Textbox 2.** Representative words of selected topics obtained from 3 groups of tweets.

Weight loss	<ul style="list-style-type: none"> <li>• Weight, lose, lost, gain, lb</li> </ul>
Eating disorder symptoms	<ul style="list-style-type: none"> <li>• Purge, binge, crave, buffet, bathroom</li> </ul>
Food and drink	<ul style="list-style-type: none"> <li>• Coke, breakfast, sandwich, chicken, yogurt</li> </ul>
Body image	<ul style="list-style-type: none"> <li>• Collar bone, thigh, fat, mirror</li> </ul>
Media or advertising or portrayals	<ul style="list-style-type: none"> <li>• Instagram, twitter, media, social, tumblr</li> </ul>
Mental illness	<ul style="list-style-type: none"> <li>• Mental, ill, breakdown, think, disorder</li> </ul>
Negative consequences	<ul style="list-style-type: none"> <li>• Sleep, hunger, stress, escape, pain</li> </ul>
Negative emotions	<ul style="list-style-type: none"> <li>• Depress, apart, alone, sad, pointless</li> </ul>
Education or awareness or treatment	<ul style="list-style-type: none"> <li>• Therapist, session, save, visit, came</li> </ul>
Recovery	<ul style="list-style-type: none"> <li>• Battle, strength, courage, stronger, inspiring</li> </ul>

## Topics Evaluation

The CorEx model results of the 3 experimental groups were reviewed and analyzed by a domain expert (AH). For group 1, which used manually identified *ED-laypeople* tweets, 54 of 70 topics were identified as meaningful, and each of them was assigned a topic theme. Similar themes were further grouped into 15 higher-level categories. The top 10 relevant tweets of each topic were reviewed and judged whether they were coherent with the summarized topic theme, and the coherence rate was calculated. [Table 2](#) lists the summary of group 1, including the identified higher-level topic categories, the number of topic themes under each category, some representative topic themes, and the coherence rates for each topic category.

For the second group that used *ED-laypeople* tweets identified by the developed classifier, 63 of 70 topics were identified as meaningful topics and were assigned topic themes. The 63 topics were further merged into 19 categories. [Table 3](#) shows a

summary of the topics in the second group, including the identified topic categories, the number of topic themes under each category, example of representative topic themes, and the coherence rates for each category.

For the third group, which used manually identified *ED-irrelevant* tweets posted by the potential ED users, 47 of 80 topics were reviewed as significant topics. The 47 topics were further merged into 19 categories. [Table 4](#) shows a summary of the results from group 3.

Compared with our previous study [17], several new topics were identified, including *Questions or Concerns*, *Reflection or Planning*, *Comorbidity*, *Ambivalence*, *Insults*, and *Diagnostic criteria*. [Textbox 3](#) shows these topic themes and example tweets.

We also explored the ED-irrelevant tweets posted by the potential ED users (ie, group 3). Selected topics and relevant tweets are listed in [Textbox 4](#).

**Table 2.** A summary of the topics using group 1 tweets (ie, manually identified ED-laypeople tweets).

Topic category	Population, n	Number of topics under each category	Representative topic themes	Coherence rate, n (%)
ED <sup>a</sup> recovery	90	9	Learning from the past; Hope; Moving forward	69 (77)
ED symptoms	70	7	Weight loss and gain; Binge-eating and purging	61 (87)
Education or awareness or treatment	60	6	ED education; ED treatment	51 (85)
Random words	60	6	Love; Big; Life; Rock	48 (80)
Negative consequences	50	5	Health damage; Feeling trapped	36 (72)
Body image	50	5	Collar bones; Thinness	39 (78)
Food and drink	30	3	Food and drink	23 (77)
Pro-ana	30	3	Pro-ana	24 (80)
Negative emotions	30	3	Guilt and shame; Fear; Sadness	23 (77)
Media or advertising or portrayals	20	2	Media and advertising	12 (60)
Comorbidity	10	1	Comorbidity	10 (100)
Reflection or planning	10	1	Reflection or planning	9 (90)
Ambivalence	10	1	Ambivalence	8 (80)
Diagnostic criteria	10	1	Diagnostic criteria	7 (70)
Questions or concerns	10	1	Questions or concerns	10 (100)

<sup>a</sup>ED: eating disorder.

**Table 3.** A summary of the topics using group 2 tweets (ie, ED-laypeople tweets identified by the developed classifiers).

Topic category	Population, n	Number of topics under each category	Representative topic themes	Coherence rate, n (%)
ED <sup>a</sup> symptoms	100	10	Restriction; Appetite suppression; Binge-eating; Purging	65 (65.0)
Education or awareness or treatment	90	9	ED education; Support group	76 (84.4)
Media or advertising or portrayals	90	9	Media or advertising	65 (72.2)
ED recovery	80	8	Passion; Hope; Love	63 (78.8)
Negative consequences	70	7	Health damage; Social	44 (62.9)
Food and drink	20	2	Food and drink	13 (65.0)
Social media	20	2	Twitter; Social media	11 (55.0)
Pro-ana	20	2	Pro-ana	16 (80.0)
Insults	20	2	Insults	18 (90.0)
Reflection or planning	20	2	Reflection or planning	17 (85.0)
Comorbidity	20	2	Comorbidity	19 (95.0)
Mental illness	10	1	Mental illness	9 (90.0)
Negative emotions	10	1	Negative emotions	8 (80.0)
Body image	10	1	Body image	8 (80.0)
Weight extremes	10	1	Weight extremes	5 (50.0)
Negative social reactions	10	1	Negative social reactions	8 (80.0)
Diagnosis	10	1	Diagnosis	9 (90.0)
Questions or concerns	10	1	Questions or concerns	6 (60.0)
Random words	10	1	Anger	9 (90.0)

<sup>a</sup>ED: eating disorder.



**Table 4.** A summary of the topics using group 3 tweets (ie, manually identified ED-irrelevant tweets posted by the potential ED users).

Topic category	Population, n	Number of topics under each category	Representative topic themes	Coherence rate, n (%)
Negative emotions or attitude	80	8	Negative emotion; Pressure; Hate	68 (85)
ED <sup>a</sup> behaviors	70	7	Restriction; Purging; Laxatives	55 (79)
Body image	50	5	Hair; Appearance	36 (72)
Exercise	40	4	Exercise	26 (65)
Media	30	3	Entertainment; Music	20 (67)
Self-harm	20	2	Self-harm	17 (85)
Negative consequences	20	2	Mental health damage	19 (95)
Communication	20	2	Seeking communication; Connection	12 (60)
Weight loss	20	2	Weight loss; Concern about weight	18 (90)
Positive emotions	20	2	Positive emotions; Encouraging	14 (70)
Holidays	20	2	Christmas; Halloween	15 (75)
Food and drink	10	1	Food and drink	9 (90)
Social media	10	1	Social media	8 (80)
Suicide	10	1	Desire for suicide	9 (90)
Appreciation	10	1	Appreciation	10 (100)
Sleep deprivation	10	1	Sleep deprivation	10 (100)
Grocery or shopping	10	1	Grocery or shopping	6 (60)
Intimate relationships	10	1	Intimate relationships	7 (70)
School	10	1	Negative social reactions	6 (60)

<sup>a</sup>ED: eating disorder.

**Textbox 3.** New topics identified from ED-laypeople tweets in groups 1 and 2.

#### Comorbidity

- “#mentalhealthawareness We're hiring. I myself suffer with 6 mental health illnesses. Anxiety, Depression, OCD, BPD, Anorexia and PTSD”
- “Got my full diagnosis list. I have major depression with psychotic features, GAD, anorexia nervosa, binge eating, purging type, ptsd, and bpd”
- “Literally my everyday life, so many diagnosis, ocd, did, bulimia, anorexia ....anxiety depression until 11 years later “bpd” or eupd as they call it in uk, its hard to tell ppl more should be done to raise awareness so much praise for @xxxxxxx”

#### Reflection or planning

- “A few months ago I was hiding in my dorm room severely depressed and relapsing from anorexia; tomorrow morning marks one month of me taking recovery seriously; is also my move-in day at a new school & I am so excited to get healthy again and for a new start and I am truly happy”
- “So yesterday was a struggle with Candy it was all around constantly and I caved . :( but I will get back on track today. I will pretend to be sick on thanksgiving and on Christmas Eve that way I don't get tempted by candy .... I'm done with looking at that scale go up. #proana”
- “Live for today, let yesterday go, and keep smiling for tomorrow. #Anorexia #ED #life #strength #dontgiveup”

#### Insults

- “Net of 0 today. My mom made me eat a lean cuisine :( she's yelling all day that I'm “anorexic”. Stfu, leave me alone I'm FAT.”

#### Diagnostic criteria

- “@xxxxxxx even at my sickest, I didn't meet the weight criteria for anorexia diagnosis. But I had severe muscle wastage at that point”
- “@xxxxxxx what a sweetheart you are. I was quoting the medically accepted diagnostic criteria for anorexia nervosa... thx for the information.”

#### Ambivalence

- “@xxxxxxx @xxxxx Hi I also have an eating disorder, and I often find it comforting. Please don't speak on behalf of every single person with an ED. We are all very different, and you never know what helps. Let's try building each other up.”
- “Worst feeling ever is missing my thin anorexia body even though logically I know it was unhealthy and killing me. And now I have to struggle everyday to live in this “fat” body. Just completely unbearable at times.”
- “My eating disorder is my worst enemy, yet my closest friend.”

**Textbox 4.** The selected topics and the representative tweets in group 3.

Negative emotions or attitudes	<ul style="list-style-type: none"> <li>• “There was a certain weight threshold I never wanted to go above. And now I’m above that. I feel like a complete utter failure.”</li> <li>• “Honestly I’m disgusting. How can I even contemplate eating when I’m this morbidly obese. Jesus.”</li> <li>• “Sorry for being bitchy guys. I’m really tired, I need a shower, I’m worried, and I’m hungry. I’m just grumpy.”</li> </ul>
Eating disorder symptoms	<ul style="list-style-type: none"> <li>• “Haven’t eaten in 25hours, will eat in 30minutes. Salad &amp; Chicken.”</li> <li>• “I haven’t properly fasted for such a long time, only restricted or binged. My fast began 3 hours ago and I’m actually excited! #fasting”</li> <li>• “@xxxxx I’m liquid fasting for at least 24hrs now.”</li> </ul>
Self-harm	<ul style="list-style-type: none"> <li>• “My wrist n cuts are getting dry, they disappear so fast”</li> <li>• “I have so many cuts at the moment. I went a bit mental last week. Everything’s just falling apart.”</li> <li>• “All because I’m so emotional, I’m gonna be flaunting some major self harming. The county fair is next week... Everyone will see...”</li> </ul>
Negative consequences	<ul style="list-style-type: none"> <li>• “I feel so empty, I just want to cut my wrist and bleed out #depressed #depressedgirl #nomoreburns #numb”</li> <li>• “crying. life and mind are falling apart again.”</li> <li>• “The self hatred is so mentally and physically draining.”</li> <li>• “I feel so mentally unstable at the moment. Like I’m on the verge of a complete mental breakdown any minute.”</li> </ul>
Communication	<ul style="list-style-type: none"> <li>• “Someone talk to me please. I’m just going mad please!”</li> <li>• “Someone please talk me out of the Popsicle I want so bad. #HelpMe”</li> <li>• “I could use someone to talk to. By talk, I mean text lol feelin kinda lonely all alone. :( #miasisters”</li> </ul>
Suicide	<ul style="list-style-type: none"> <li>• “I’m getting recurrent suicidal thoughts. My mental health is just awful at the moment. I feel very self destructive and unstable. I hate it.”</li> <li>• “I want to die so bad but having to commit suicide will change my family’s life and other peoples outlook on them”</li> <li>• “I’m so depressed, I feel like the only escape is suicide. Life is so pointless, you live in mental hell, you live with physical pain, you end up alone hen death.”</li> </ul>
Sleep deprivation	<ul style="list-style-type: none"> <li>• “It’s 4:40am and I can’t fall asleep. I’m yawning cause I’m tired, but I can’t find sleep. I have too many things going through my mind.”</li> <li>• “Laying in bed and my stomach won’t stop churning, think I’m gonna be having a restless night tonight!”</li> <li>• “In the past two days I’ve drank 23 cups of coffee and actually hallucinated. I’m not sure if its from sleep deprivation or caffeine #Whoops”</li> </ul>
Laxative usage	<ul style="list-style-type: none"> <li>• “Uh-oh! Can start to feel the lax cramps now.....”</li> <li>• “I can start to feel my laxatives kicking in. I took 5 so far. Just hate the cramps when I take laxatives. :-/”</li> <li>• “I’ve been using a lot of laxatives lately. It doesn’t seem to do much except give me cramps.”</li> </ul>

## Discussion

### Identification of ED-Relevant Tweets and Topic Modeling

In earlier studies, the identification of ED-relevant tweets was based on a manual search method using ED keywords or hashtags and filtered by manually curated rules [19,21]. This

approach was not efficient and may have included numerous irrelevant tweets owing to the ambiguity of the keywords. In this study, we developed deep learning classifiers to automatically identify relevant tweets posted by *ED-laypeople*. We simplified the classification task into a 2-step process and achieved good performance with 2 supervised deep learning classifiers (ie, CNN for step 1 and LSTM for step 2) and

obtained reasonable results with  $F_1$  scores of 0.89 and 0.90, respectively. For content analysis of the identified tweets, we applied topic modeling, an unsupervised method, and manual review, which could summarize comprehensive information from tens of thousands of tweets. This approach is more efficient compared with the completely manual review, which could only cover several hundreds of tweets [21].

### Analysis of Topic Modeling Results

The overall coherence rate for the 3 experimental groups was 77.07% (1264/1640), which indicates high quality of the topics produced by the CorEx model. The categories and higher-level themes of the produced topics were summarized by the domain expert (AH). Some of these categories may seem unclear or disparate if not considered within the right context. For example, the *Insults* category referred to using ED symptoms or terms to insult someone (eg, using the word *anorexic* as a derogatory term). Therefore, this content was ED-related, although not pertaining to a traditional ED topic (eg, ED symptoms). This principle was similar for the other categories. For instance, the category of *Questions* referred to a range of ED content (eg, “@instagram This is not a difficult question - how do I escalate a complaint about dangerous, self-harm-encouraging, pro ana content that your algorithm has deemed safe?”), although the organizing principle for this group of Tweets was question-asking (ie, all tweets were posed in the form of a question). We believe that the diversity of topics identified by this algorithm is a strength of this investigation. Producing a range of topics will allow researchers to better understand the social media content on EDs, including content that would not have been expected. This information could ultimately inform future mechanistic and treatment research on EDs. For instance, we would not have anticipated that ED terminology would be used as insults on social media. We hypothesize that this type of language could further stigmatize this set of disorders, which present a target for future prevention efforts. In addition, the *Questions* category identified issues important to potential ED users on social media, such as the need for better social media monitoring and blocking of content that could be detrimental to individuals with, or vulnerable to developing, an ED. This further highlights the importance of generating algorithms for identifying ED-related content on social media. Improvement of these methods could be used by social media platforms such as Twitter to improve filtering practices for harmful content and/or to provide appropriate mental health resources to individuals posting or viewing such content.

The first and second groups focused on the *ED-laypeople* tweets posted by manually identified potential ED users. One aim of comparing these 2 groups was to verify that the *ED-laypeople* tweets identified by manually specified rules and a machine learning classifier could produce similar topics. The 2 groups were found to have 14 overlapping topic categories, such as *Negative consequences*, *ED symptoms*, *Education or awareness or treatment*, and *ED recovery*. This result was in agreement with our previous study [17]. There were slight discrepancies in the identified categories between the 2 groups. *Body image* uniquely appeared among the highest frequency topics in group 1, whereas *Media or advertising or portrayals* uniquely appeared among the highest frequency topics in group 2. The higher

prevalence of the *Media or advertising or portrayals* topic category in group 2 reveals a pitfall of the developed machine learning classifier; some *ED-promotional and education* tweets were misclassified as *ED-laypeople* tweets by the classifier, leading to a larger *Media or advertising or portrayals* topic category. When we manually identified the *ED-laypeople* tweets, we could check both the content of the tweets and the user profile of the accounts that posted the tweets. However, the developed classifier only used the content from the tweets themselves and did not incorporate profile information. Including user information such as usernames as features in the classifier would have the potential to mitigate this misclassification problem. Furthermore, group 1 had one unique topic category, *Ambivalence*, occupying 2% (1/54) of total meaningful topics in this group, whereas group 2 had unique categories of *Negative social reactions*, *Insults*, *Mental illness*, and *Social media*, occupying 10% (6/63) of the total meaningful topics. In general, the 2 groups of tweets produced highly similar topics, indicating that our machine learning classifier was mostly as effective as the manual method. This is meaningful because it suggests that this method could be used to identify ED-relevant social media content for future larger scale investigations. In addition, such methods could ultimately aid in identifying at-risk groups for whom prevention efforts could be targeted, which is especially important given the low level of ED detection in typical practice [2].

According to [Textbox 4](#), 6 new topics were identified compared with our previous study [17]. These topics are consistent with earlier literature on EDs and may provide novel insights into the thoughts and experiences of individuals with EDs. Several areas of the content reflected topics that may be relevant to understanding the decisional mechanisms involved in these disorders. The *Reflection or planning* category (eg, “A few months ago I was hiding in my dorm room severely depressed and relapsing from anorexia; tomorrow morning marks one month of me taking recovery seriously; is also my move-in day at a new school.”) and *Ambivalence* category (eg, “My eating disorder is my worst enemy, yet my closest friend.”) reflect the strong pull of individuals with EDs toward and against ED symptoms [26]. This suggests that for some, the decision about whether to engage in ED behavior involves consideration of the pros and cons of engaging in these behaviors and planning future actions [26]. In future research, it would be informative to determine whether individuals producing tweets reflecting the deliberative processes of weighing pros and cons of ED behaviors vary in a clinically relevant manner from those not producing these tweets, as there is a suggestion that deliberative processes might characterize earlier stages of illness [18]. In addition, the *Comorbidity* and *Diagnostic criteria* categories highlight the importance of considering heterogeneity across a range of severity and symptom profiles in ED research and treatment [27]. Many ED researchers have highlighted the importance of considering certain ED characteristics (eg, whether a person is emotionally dysregulated vs constrained) in planning treatments [28]. In addition, it has long been acknowledged that *Diagnostic criteria* categories fail to capture many individuals who do not present with classic ED symptoms (eg, meet all criteria for anorexia nervosa but are not underweight) [29]. These topics also might characterize content

from clinically unique subgroups of individuals with EDs, warranting further consideration.

The summarized topics for the group 3 tweets are listed in [Textbox 3](#). The high-frequency topics ( $n > 5$ ) identified from the ED-irrelevant tweets included *Negative emotions or attitudes* ( $n=8$ ), *Eating disorder behaviors* ( $n=7$ ), and *Body image* ( $n=5$ ). The content identified in these analyses reveals what other topics are common to potential ED users in their daily lives. One notable feature of these results is that the content of many of the tweets identified as *ED-irrelevant* pertains explicitly to ED cognitive and behavior symptoms (eg, *ED symptoms* and *Body image* categories). One interpretation is that for individuals with ED, the experience with their disorder becomes so pervasive that it infiltrates content that is not explicitly intended as ED-relevant. However, this could also indicate that the algorithm needs to be further refined to capture all the content relevant to EDs. An additional finding from this analysis is that many of the tweets demonstrate that negative emotion and pain predominate the experience of having an ED, as reflected in the categories *Negative emotions or attitude*, *Negative consequences*, *Self-harm*, and *Suicide*. These categories correspond with theories that ED symptoms may result from a surfeit of negative emotions and that the symptoms themselves may function to alleviate emotional pain in a similar fashion to self-harm and suicide planning [30-32].

### Limitations

We developed machine learning classifiers that could identify ED-relevant tweets with high performance, but there is still a

small percentage of misclassified tweets, especially in the task of differentiating the *ED-promotional and education* versus *ED-laypeople* tweets. This may be partially due to the short length of some tweets, which makes them difficult to classify. Furthermore, these 2 types of tweets are sometimes semantically similar. With the increase in the number of collected tweets, there will be a large number of misclassified tweets, although the misclassification rate is low, which will further influence the results of topic modeling as it will produce some noise topics. Another limitation is that we cannot collect all the tweets in the entire timeline of our target potential ED users owing to the restriction of Twitter API, which may lead to other useful topics being missed.

### Conclusions

Our study developed a 2-step process using 2 classifiers (ie, CNN and LSTM) that could automatically identify ED-relevant tweets posted by the potential ED users. The  $F_1$  scores of the 2 classifiers were 0.89 and 0.90, respectively. A CorEx model was applied on the tweets identified by the classifiers and those identified by a traditional manual method separately. Highly overlapping topics were produced. Through a review of these topics by a domain expert, important features of the social media content of potential ED were identified. These findings provided novel insights into the experience of having an ED, which could be expanded upon in future research using the methods derived in this investigation.

### Acknowledgments

This study was supported by the National Center for Complementary and Integrative Health of the National Institutes of Health (NIH) under Award Number R01AT009457 (principal investigator [PI]: RZ), the National Institute of Mental Health under Award Number K23 MH112867 (PI: AH), and the National Science Foundation (NSF) under Award Number 1734134 (PI: JB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the NSF.

### Conflicts of Interest

None declared.

### References

1. Hudson JI, Hiripi E, Pope HG, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 2007 Feb 01;61(3):348-358 [[FREE Full text](#)] [doi: [10.1016/j.biopsych.2006.03.040](https://doi.org/10.1016/j.biopsych.2006.03.040)] [Medline: [16815322](#)]
2. Kutz AM, Marsh AG, Gunderson CG, Maguen S, Masheb RM. Eating Disorder Screening: a Systematic Review and Meta-analysis of Diagnostic Test Characteristics of the SCOFF. *J Gen Intern Med* 2020 Mar;35(3):885-893. [doi: [10.1007/s11606-019-05478-6](https://doi.org/10.1007/s11606-019-05478-6)] [Medline: [31705473](#)]
3. Eddy KT, Tabri N, Thomas JJ, Murray HB, Keshaviah A, Hastings E, et al. Recovery From Anorexia Nervosa and Bulimia Nervosa at 22-Year Follow-Up. *J Clin Psychiatry* 2017 Feb;78(2):184-189. [doi: [10.4088/JCP.15m10393](https://doi.org/10.4088/JCP.15m10393)] [Medline: [28002660](#)]
4. Steinhausen H. The outcome of anorexia nervosa in the 20th century. *Am J Psychiatry* 2002 Aug;159(8):1284-1293. [doi: [10.1176/appi.ajp.159.8.1284](https://doi.org/10.1176/appi.ajp.159.8.1284)] [Medline: [12153817](#)]
5. Bauer S, Kindermann SS, Moessner M. [Prevention of eating disorder: a review]. *Z Kinder Jugendpsychiatr Psychother* 2017 Sep;45(5):403-413. [doi: [10.1024/1422-4917/a000506](https://doi.org/10.1024/1422-4917/a000506)] [Medline: [27951744](#)]
6. Berkman ND, Bulik CM, Brownley KA, Lohr KN, Sedway JA, Rooks A, et al. Management of eating disorders. *Evid Rep Technol Assess (Full Rep)* 2006 Apr(135):1-166. [Medline: [17628126](#)]
7. Kass AE, Kolko RP, Wilfley DE. Psychological treatments for eating disorders. *Curr Opin Psychiatry* 2013 Nov;26(6):549-555 [[FREE Full text](#)] [doi: [10.1097/YCO.0b013e328365a30e](https://doi.org/10.1097/YCO.0b013e328365a30e)] [Medline: [24060917](#)]

8. Paul MJ, Dredze M. You are what you tweet: Analyzing Twitter for public health. 2011 Presented at: Fifth International AAAI Conference on Weblogs and Social Media; 17-21, July, 2011; Barcelona, Spain p. A.
9. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging Social Media to Promote Public Health Knowledge: Example of Cancer Awareness via Twitter. *JMIR Public Health Surveill* 2016;2(1):e17 [FREE Full text] [doi: [10.2196/publichealth.5205](https://doi.org/10.2196/publichealth.5205)] [Medline: [27227152](https://pubmed.ncbi.nlm.nih.gov/27227152/)]
10. Musaev A, Britt RK, Hayes J, Britt BC, Maddox J, Sheinidashtegol P. Study of Twitter communications on cardiovascular disease by state health departments. 2019 Presented at: InInternational Conference on Web Services (pp. ). Springer, Cham; 2019 Jun 25; Milan, Italy p. 181-189. [doi: [10.1007/978-3-030-23499-7\\_12](https://doi.org/10.1007/978-3-030-23499-7_12)]
11. Bian J, Topaloglu U, Yu F. Towards large-scale twitter mining for drug-related adverse events. 2012 Oct Presented at: InProceedings of the 2012 international workshop on Smart health and wellbeing (pp. ); 2012 Oct 29; Maui, USA p. 25-32. [doi: [10.1145/2389707.2389713](https://doi.org/10.1145/2389707.2389713)]
12. Zhang H, Wheldon C, Dunn AG, Tao C, Huo J, Zhang R, et al. Mining Twitter to assess the determinants of health behavior toward human papillomavirus vaccination in the United States. *J Am Med Inform Assoc* 2020 Feb 01;27(2):225-235. [doi: [10.1093/jamia/ocz191](https://doi.org/10.1093/jamia/ocz191)] [Medline: [31711186](https://pubmed.ncbi.nlm.nih.gov/31711186/)]
13. Modave F, Zhao Y, Krieger J, He Z, Guo Y, Huo J, et al. Understanding Perceptions and Attitudes in Breast Cancer Discussions on Twitter. *Stud Health Technol Inform* 2019 Aug 21;264:1293-1297 [FREE Full text] [doi: [10.3233/SHTI190435](https://doi.org/10.3233/SHTI190435)] [Medline: [31438134](https://pubmed.ncbi.nlm.nih.gov/31438134/)]
14. Zhao Y, Guo Y, He X, Wu Y, Yang X, Prospero M, et al. Assessing mental health signals among sexual and gender minorities using Twitter data. *Health Informatics J* 2020 Jun;26(2):765-786. [doi: [10.1177/1460458219839621](https://doi.org/10.1177/1460458219839621)] [Medline: [30969146](https://pubmed.ncbi.nlm.nih.gov/30969146/)]
15. Wang Y, Zhao Y, Zhang J, Bian J, Zhang R. Detecting associations between dietary supplement intake and sentiments within mental disorder tweets. *Health Informatics J* 2020 Jun;26(2):803-815. [doi: [10.1177/1460458219867231](https://doi.org/10.1177/1460458219867231)] [Medline: [31566452](https://pubmed.ncbi.nlm.nih.gov/31566452/)]
16. Hicks A, Hogan WR, Rutherford M, Malin B, Xie M, Fellbaum C, et al. Mining Twitter as a First Step toward Assessing the Adequacy of Gender Identification Terms on Intake Forms. *AMIA Annu Symp Proc* 2015;2015:611-620 [FREE Full text] [Medline: [26958196](https://pubmed.ncbi.nlm.nih.gov/26958196/)]
17. Zhou S, Bian J, Zhao Y, Haynos AF, Rizvi R, Zhang R. Analysis of Twitter to Identify Topics Related to Eating Disorder Symptoms. *IEEE Int Conf Healthc Inform* 2019 Jun;2019 [FREE Full text] [doi: [10.1109/ichi.2019.8904863](https://doi.org/10.1109/ichi.2019.8904863)] [Medline: [32030368](https://pubmed.ncbi.nlm.nih.gov/32030368/)]
18. Walsh BT. The enigmatic persistence of anorexia nervosa. *Am J Psychiatry* 2013 May;170(5):477-484 [FREE Full text] [doi: [10.1176/appi.ajp.2012.12081074](https://doi.org/10.1176/appi.ajp.2012.12081074)] [Medline: [23429750](https://pubmed.ncbi.nlm.nih.gov/23429750/)]
19. Arseniev-Koehler A, Lee H, McCormick T, Moreno MA. #Proana: Pro-Eating Disorder Socialization on Twitter. *J Adolesc Health* 2016 Jun;58(6):659-664. [doi: [10.1016/j.jadohealth.2016.02.012](https://doi.org/10.1016/j.jadohealth.2016.02.012)] [Medline: [27080731](https://pubmed.ncbi.nlm.nih.gov/27080731/)]
20. Kenny TE, Boyle SL, Lewis SP. #recovery: Understanding recovery from the lens of recovery-focused blogs posted by individuals with lived experience. *Int J Eat Disord* 2020 Aug;53(8):1234-1243. [doi: [10.1002/eat.23221](https://doi.org/10.1002/eat.23221)] [Medline: [31886573](https://pubmed.ncbi.nlm.nih.gov/31886573/)]
21. Cavazos-Rehg PA, Krauss MJ, Costello SJ, Kaiser N, Cahn ES, Fitzsimmons-Craft EE, et al. "I just want to be skinny": A content analysis of tweets expressing eating disorder symptoms. *PLoS One* 2019;14(1):e0207506 [FREE Full text] [doi: [10.1371/journal.pone.0207506](https://doi.org/10.1371/journal.pone.0207506)] [Medline: [30650072](https://pubmed.ncbi.nlm.nih.gov/30650072/)]
22. Alghamdi R, Alfalqi K. A Survey of Topic Modeling in Text Mining. *Int. J. Adv. Comput. Sci. Appl* 2015;6(1). [doi: [10.14569/IJACSA.2015.060121](https://doi.org/10.14569/IJACSA.2015.060121)]
23. Barde BV, Bainwad AM. An overview of topic modeling methods and tools. 2017 Jun Presented at: International Conference on Intelligent Computing and Control Systems (ICICCS); 2017 Jun 15; Madurai, India p. 745-750. [doi: [10.1109/iccons.2017.8250563](https://doi.org/10.1109/iccons.2017.8250563)]
24. Yan X, Guo J, Lan Y, Cheng X. A bitern topic model for short texts. 2013 May Presented at: Proceedings of the 22nd international conference on World Wide Web; 2013 May 13; Rio de Janeiro, Brazil p. 1445-1456. [doi: [10.1145/2488388.2488514](https://doi.org/10.1145/2488388.2488514)]
25. Gallagher RJ, Reing K, Kale D, Ver Steeg G. Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge. *Transactions of the Association for Computational Linguistics* 2017 Dec;5:529-542. [doi: [10.1162/tacl\\_a\\_00078](https://doi.org/10.1162/tacl_a_00078)]
26. Serpell L, Treasure J, Teasdale J, Sullivan V. Anorexia nervosa: friend or foe? *Int J Eat Disord* 1999 Mar;25(2):177-186. [doi: [10.1002/\(sici\)1098-108x\(199903\)25:2<177::aid-eat7>3.0.co;2-d](https://doi.org/10.1002/(sici)1098-108x(199903)25:2<177::aid-eat7>3.0.co;2-d)] [Medline: [10065395](https://pubmed.ncbi.nlm.nih.gov/10065395/)]
27. Wildes JE, Marcus MD. Alternative methods of classifying eating disorders: models incorporating comorbid psychopathology and associated features. *Clin Psychol Rev* 2013 Apr;33(3):383-394 [FREE Full text] [doi: [10.1016/j.cpr.2013.01.006](https://doi.org/10.1016/j.cpr.2013.01.006)] [Medline: [23416343](https://pubmed.ncbi.nlm.nih.gov/23416343/)]
28. Wildes JE, Marcus MD, Crosby RD, Ringham RM, Dapelo MM, Gaskill JA, et al. The clinical utility of personality subtypes in patients with anorexia nervosa. *J Consult Clin Psychol* 2011 Oct;79(5):665-674 [FREE Full text] [doi: [10.1037/a0024597](https://doi.org/10.1037/a0024597)] [Medline: [21767000](https://pubmed.ncbi.nlm.nih.gov/21767000/)]
29. Dunn EC, Geller J, Brown KE, Bates ME. Addressing the EDNOS issue and improving upon the utility of DSM-IV: classifying eating disorders using symptom profiles. *Eur Eat Disord Rev* 2010;18(4):271-280. [doi: [10.1002/erv.1005](https://doi.org/10.1002/erv.1005)] [Medline: [20552559](https://pubmed.ncbi.nlm.nih.gov/20552559/)]

30. Haynos AF, Fruzzetti AE. Anorexia nervosa as a disorder of emotion dysregulation: evidence and treatment implications. *Clin Psychol Sci Prac* 2011;18(3):183-202. [doi: [10.1111/j.1468-2850.2011.01250.x](https://doi.org/10.1111/j.1468-2850.2011.01250.x)]
31. Pisetsky EM, Haynos AF, Lavender JM, Crow SJ, Peterson CB. Associations between emotion regulation difficulties, eating disorder symptoms, non-suicidal self-injury, and suicide attempts in a heterogeneous eating disorder sample. *Compr Psychiatry* 2017 Feb;73:143-150 [FREE Full text] [doi: [10.1016/j.comppsy.2016.11.012](https://doi.org/10.1016/j.comppsy.2016.11.012)] [Medline: [27978502](https://pubmed.ncbi.nlm.nih.gov/27978502/)]
32. Wang SB, Borders A. The unique effects of angry and depressive rumination on eating-disorder psychopathology and the mediating role of impulsivity. *Eat Behav* 2018 Apr;29:41-47. [doi: [10.1016/j.eatbeh.2018.02.004](https://doi.org/10.1016/j.eatbeh.2018.02.004)] [Medline: [29477016](https://pubmed.ncbi.nlm.nih.gov/29477016/)]

## Abbreviations

**API:** application programming interface  
**BTM:** Biterm Topic Model  
**CNN:** convolutional neural network  
**CorEx:** Correlation Explanation  
**ED:** eating disorder  
**GB:** gradient boosting trees  
**LDA:** latent Dirichlet allocation  
**LSTM:** long short-term memory  
**NB:** naïve Bayes  
**NIH:** National Institutes of Health  
**NSF:** National Science Foundation  
**PI:** principal investigator  
**RF:** random forest  
**SVM:** support vector machine

*Edited by C Lovis; submitted 15.02.20; peer-reviewed by L Cui, R Rodgers, K Reuter; comments to author 28.04.20; revised version received 14.07.20; accepted 06.09.20; published 30.10.20*

*Please cite as:*

Zhou S, Zhao Y, Bian J, Haynos AF, Zhang R  
*Exploring Eating Disorder Topics on Twitter: Machine Learning Approach*  
*JMIR Med Inform* 2020;8(10):e18273  
URL: <http://medinform.jmir.org/2020/10/e18273/>  
doi: [10.2196/18273](https://doi.org/10.2196/18273)  
PMID: [33124997](https://pubmed.ncbi.nlm.nih.gov/33124997/)

©Sicheng Zhou, Yunpeng Zhao, Jiang Bian, Ann F Haynos, Rui Zhang. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 30.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.