Jason Ho (jasonho4)

CS410

Technology Review on Applications of LDA in Social Media

With the abundance of user data on social media in today's world, there is a lot to be learned through analysis of text data displayed on these platforms. People go on social media to share various ideas from personal experiences and thoughts to advertisements of various products. Each of these unique posts comes with an assortment of emotions and much of these emotions can be found in the text body of the post through topic modeling techniques. Recent studies have utilized the topic modeling technique, Latent Dirichlet Allocation or LDA, to identify patterns in the text that display anxiety and depression. This is useful in that it helps build a predictor to identify individuals who may be suffering from mental health issues. This technology review will cover and compare research on applying LDA to social media platforms detailed in Beyond *LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter* by Philip Resnik2,4 , William Armstrong1,4 , Leonardo Claudino1,4 , Thang Nguyen3 , Viet-An Nguyen1,4 , and Jordan Boyd-Graber3, *Detecting anxiety on Reddit* by Shen and Rudzicz, and *Facebook language predicts depression in medical records* by Eichstaedt, Smith, Merchant, Ungar, Crutchley, Preoţiuc-Pietro, Asch, Schwartz. Each of these papers go into utilizing LDA on different social media platforms Twitter, Reddit, and Facebook.

In Beyond *LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter*, LDA and sLDA, or supervised LDA, was applied on about 3 million tweets across 2000 unique Twitter users. In the base unsupervised LDA, the model was empirically found to work best with 50 topics. These topics were hand-reviewed by a licensed clinical psychologist to ensure the topics are relevant to depression. This shows that domain knowledge is important even when applying an unsupervised model and as described by the paper, is a “poor-man’s version of interactive topic modeling”. Utilizing this domain knowledge is especially useful because the topics generated from the unsupervised LDA serve as informative priors that are fed into the supervised LDA model. That is, the expert in the field labels the topics found from unsupervised LDA with a valence of (n, p, and e) where n denotes negative valence that indicates depression, p denotes positive valence that indicates no depression, and e denotes a neutral value that does not give a strong assessment of either valence. This labeled data can then be used in supervised LDA and supervised nested LDA to identify any hints of neuroticism in the Twitter posts.

The next paper, *Detecting anxiety on Reddit,* takes a similar approach to the Twitter study by utilizing LDA on labeled data. However, the Reddit study differs in that its data contains only a binary label: the post is either anxiety related or is a control rather than the multi-class valences described in the Twitter paper. Also, the Reddit study obtains its labeled data by assuming posts coming from subreddits (sub-topic forums on Reddit) related to anxiety like r/anxiety, r/panicparty, r/healthanxiety, and r/socialanxiety automatically fall into the anxiety class. Whereas, the posts from other subreddit communities that are also posted in first-person perspective to match the writing perspective of anxiety subreddits are automatically assumed to be in the control class. This form of data collection is similar to the Twitter paper in that some domain knowledge is used but differs in that more assumptions were made compared to the expert-level knowledge in the Twitter study. Another interesting take that the Reddit paper takes that deviates from the Twiiter paper is how the Reddit paper built the LDA model with both in-domain (Reddit) and outside domain (Twitter) corpus'. Twitter was deemed similar enough of a social platform to Reddit that its training data would also be suitable for the LDA model. The results showed that building the LDA with in-domain data led to better accuracies compared to outside domain but the LDA model built from Twitter data had more "effective feature vectors". This could be due to the higher number of training data in the Twitter dataset which leads to a more complex model.

Finally, the *Facebook language predicts depression in medical records* paper goes over using LDA on Facebook posts. The study was conducted on Facebook posts from 683 patients from the same institution, 114 of which had a history of depression in their records. The usage of LDA in this paper again follows a similar pattern to what was described in the Twitter paper. 200 topics were extracted from these posts to be used as language variables in a predictive model. The study measures its models using the AUC metric. The predictive model yielded an AUC of 0.69. What separates this research from the previous two studies is how this Facebook study analyzes context. The study considered temporal context for example, citing how previous studies discovered that depressed individuals tend to post more in the evening than other times of day. Metafeatures were also analyzed (i.e. the post length and post frequency help with predicting depressed individuals). However, ultimately temporal context and metafeatures did not yield any significant improvement in predictive power. This suggests "the language content captures the depression-related variance in other feature groups" with limiting the data to Facebook posts that were posted six months prior to the depression diagnosis only yielding an improvement of an AUC of 0.72. This is interesting because it shows how the raw language variables obtained from LDA can be sufficient enough for building predictive models in certain scenarios.

All in all, it is interesting to see the different approaches to using LDA to predict depression and anxiety in individuals for different social media platforms. All of these studies share the similar component of utilizing LDA as a generator for features to be used in another predictive model but differ in how the data is formulated. The perspectives on the training data and assumptions also differed and led to several unique approaches using the same LDA technique. It would be interesting to see other or future research show LDA used to identify other emotions aside from depression and anxiety with a similar model creation procedure (using LDA to find useful features for the predictive model of choice).

Works Cited

Beyond LDA: Exploring Supervised Topic Modeling for Depression-Related Language in Twitter

Philip Resnik2,4 , William Armstrong1,4 , Leonardo Claudino1,4 , Thang Nguyen3 , Viet-An Nguyen1,4 , and Jordan Boyd-Graber3,5

Detecting anxiety on Reddit

Judy Hanwen Shen, Frank Rudzicz

Facebook language predicts depression in medical records

Johannes C. Eichstaedt, Robert J. Smith, Raina M. Merchant, Lyle H. Ungar, Patrick Crutchley, Daniel Preoţiuc-Pietro, David A. Asch, H. Andrew Schwartz