BUILDING A MEDICAL QUERY CHATBOT FOR ANXIETY DISORDERS BASED ON NATURAL LANGUAGE PROCESSING OF DSM-5 DATA

Apoorv Saraogee
Masters of Data Science Course 453
Natural Language Processing Section 56
April 3, 2023

Abstract

There is a growing shortage of manpower that can answer frequently asked questions in the medical field – both for patients and providers. Chatbots and automated assistants can leverage natural language processing techniques for useful interactions with humans and reduce the overall healthcare burden. A suitable corpus was derived from the Diagnostics and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) using the chapters related to anxiety disorders. Term frequency – inverse document frequency (TF-IDF) and a cosine similarity matrix is the classic method and is employed in this study for the vectorization of sentences and ranking of responses respectively. The logic and sensibility of responses were compared for five questions in each model. Various parameters were varied such as stop words for text preprocessing, differing corpus parameters as well as the inclusion of contextual sentences in the response. The best model utilized stop words preprocessing and the inclusion of an additional sentence for the user's contextual understanding of the chatbot's response. However, a key drawback is the requirement for using exact keywords in training chatbots with TF-IDF. Future studies would incorporate equivalent classes or embedding representations to gain deep semantic information.

Keywords: chatbot, medical, DSM-5, anxiety, NLP, TF-IDF

BUILDING A MEDICAL QUERY CHATBOT FOR ANXIETY DISORDERS BASED ON NATURAL LANGUAGE PROCESSING OF DSM-5 DATA

Chatbots are artificial intelligence powered software applications based on natural language processing techniques to interact with humans typically via text. Natural language processing has been reliably used to make query systems or chatbots with CHAT-80 being the first well known example in the field of geography (Nilsson, 2009). More recent chatbot implementations have been made for medical student assistance (Kazi et al., 2012), medical diagnosis (Lee et al., 2021) and managing neurological diseases (Ireland et al., 2016; Li, 2019). In the medical field, chatbots and automated assistant software can help address the current shortage of manpower that can answer frequently asked questions in the medical field and reduce the overall healthcare cost burden (Das et al., 2022). Indeed, chatbots are not new to the medical field with the first chatbot (ELIZA) introducing the concept of 'keyword' recognition and ranking for psychotherapy applications in the 1960s (Weizenbaum, 1966).

The impact in the mental health field is of particular focus as cost of treatment is the top concern (Center for Behavioral Health Statistics, 2020). This study focuses on anxiety disorders in the mental health field using a corpus derived from the Diagnostic and Statistical Manual for Mental Disorders, Fifth Edition (DSM-5). The common features between the human input and tokenized sentence features are analyzed using a cosine similarity matrix. The chatbot would aim to help medical professionals gain reliable medical definitions and answer commonly asked questions efficiently. Experiments will include different data wrangling methods for the vectorization of sentences from subsets of or the whole DSM-5, stop word removal and the inclusion of contextual sentences. Features of human interactions such as greetings and invalid inputs are addressed in the code. Suitability of the model is compared by evaluating sensibility of sample output conversations for five questions. Other research questions will include training time for the models.

Literature Review

While complex models have been employed with various corpus for chatbot query applications, the DSM-5 has not been used as a corpus for medical queries by itself. The DSM-5 is included in the ontologies used in the Unified Medical Language System (UMLS) (Bodenreider, 2004) which is used as a corpus for general medical queries by few others with a traditional database (Kazi et al., 2012) and a graph database (Tjokro, 2017) using UMLS. Queries provided satisfactory single sentence results in a traditional database chatbot (Kazi et al., 2012) with concept entities embedded into four constrained types. This type of complex modeling of the semantic understanding of sentences with concept modeling and surveys with medical students was required to gain meaningful interactions with humans. Graph databases required the development of the graph using domain knowledge and a symptom disease matching (SDM) algorithm (Tjokro, 2017). In the absence of survey data from mental health practitioners, the corpus in this study chose to keep it more general with a simple structure of a chatbot (Dass, 2018) and only the DSM-5 ontology focusing on the anxiety chapter. Information retrieval is documented to be a useful application of chatbots for medical providers as it quickly provides an index to relevant information from a trusted source (Kaur et al., 2021).

Methods

This study analyzes a corpus of medical literature in the mental health field known as the DSM-5 (American Psychiatric Association, 2013). This data was preprocessed into plain text format for the whole book along with a subset created for anxiety (pages 189-233). Analysis is done in a Jupyter notebook using code from a similar chatbot (Dass, 2018) and kernels for Python3 run locally with the sklearn package for vectorization with term frequency-inverse document frequency (tf-idf) and the nltk package to tokenize the words and sentences using WordNet lemmatization. Line breaks were not processed correctly inside sentences and after sentences from the pdf to plain text and were removed manually in Python3 ('\n', '- '). The user input was compared with the corpus using a cosine similarity matrix to rank sentences in the corpus. There were three chatbot models trained with descriptions shown in Table 1 below. Chatbots were also developed using the whole corpus of the DSM-5, but were omitted due to significantly longer time taken for relevant responses. Common greeting inputs, instructions to end the chatbot as well as an error message for invalid inputs were also added using while loops.

Table 1: Chatbot details with NLP architectures

Model number	**Description**
1	Corpus – stopword processing, single sentence output
2	Corpus – without stopword preprocessing, single sentence output
3	Corpus – stopword preprocessing, two sentence output

Results for each of the experiments were analyzed with the suitability of responses to five questions. These questions target areas in the ontology/knowledge structure for a DSM-5 chapter using Anxiety Disorders (page 189) as an example in Figure 1 on the next page. The questions are also reproduced below:

1. What is the prevalence of separation anxiety disorder in the United States?
2. What are the diagnostic features of panic attacks?
3. What is the essential feature of agoraphobia?
4. Does depression and anxiety have high comorbidity?
5. Is there a relationship between speaking and anxiety?

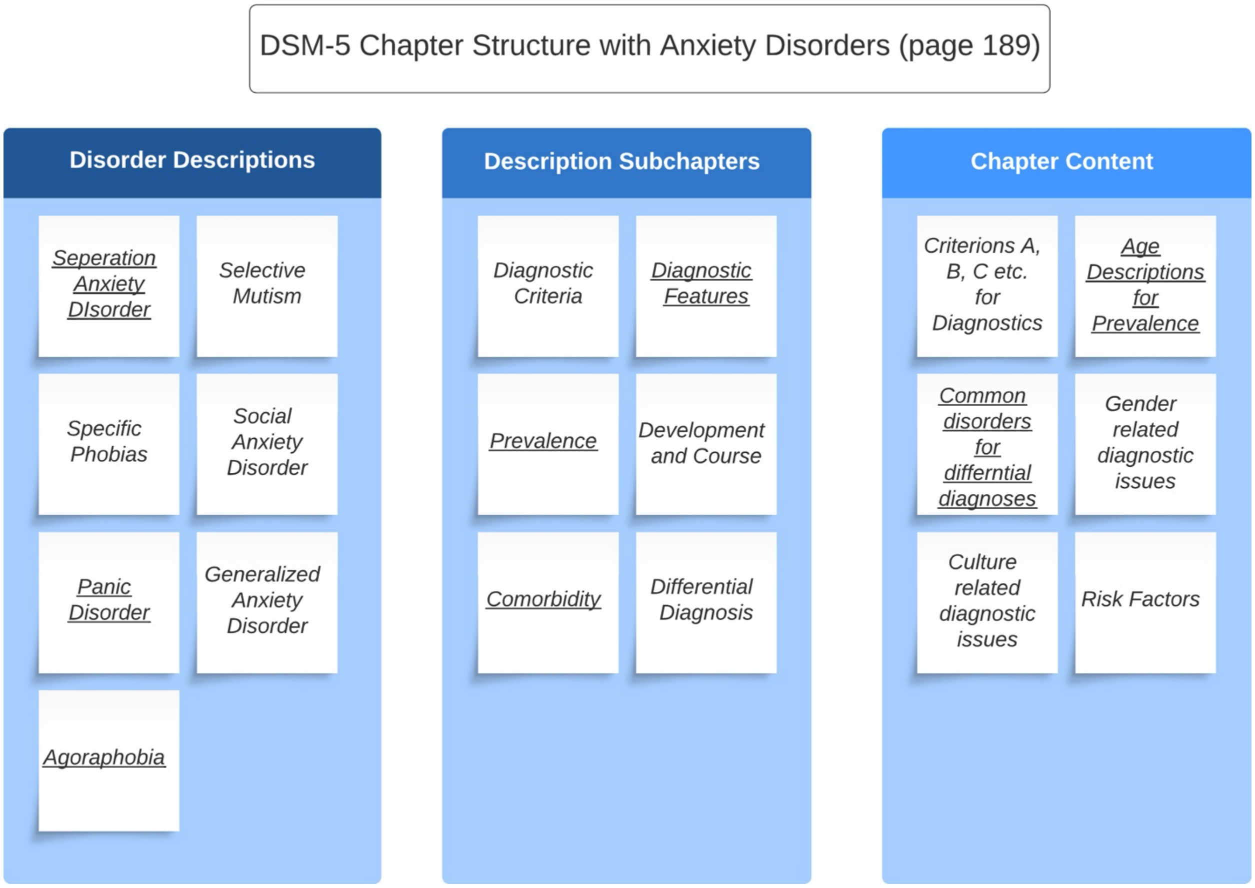

Figure 1. Ontology for chapter structures in the DSM-5. The leftmost descriptors include the main headings for each chapter within anxiety disorders. The subchapters in the middle describe the headings given for each disorder. The rightmost column includes descriptors for chapter content for disorders and subchapters. Underlined blocks indicate a relationship with questions reproduced above.

Results

Responses of the chatbot model to user questions outlined in Methods are shown in Table 2 below. The model chatbot responses are interpreted based on the suitability of the response. Model 3 with stop word removal and including an additional contextual sentence performed the best and had the most suitable responses based on the interpretation. However, some use cases may see Model 1 performing better if they require concise responses. Mostly, the additional sentence in Model 3 was useful in adding context. Model 2, which did not include preprocessing for stop word removal performed the worst.

Table 2: Chatbot model responses to questions (Q) and interpreted as best or most suitable (I)

Q	Model 1 Response	Model 2 Response	Model 3 Response	I
1	in adolescents in the united s tates, the 12-month prevalen ce is 1.6%.	in adolescents in the united s tates, the 12-month prevalen ce is 1.6%.	in adolescents in the united states, the 12 -month prevalence is 1.6%.separation an xiety disorder decreases in prevalence fro m childhood through adolescence and ad ulthood and is the most prevalent anxiety disorder in children younger than 12 year s.	3
2	features such as onset after a ge 45 years or the presence o f atypical symptoms during a panic attack (e.g., vertigo, lo ss of consciousness, loss of b ladder or bowel control, slur red speech, amnesia) sugges t the possibility that another medical condition or a subst ance may be causing the pan ic attack symptoms.other me ntal disorders with panic atta cks as an associated feature (e.g., other anxiety disorders and psychotic disorders).	panic attacks are uncommon , and unexpected panic attac ks are rare, in preadolescent children.	features such as onset after age 45 years or the presence of atypical symptoms dur ing a panic attack (e.g., vertigo, loss of c onsciousness, loss of bladder or bowel co ntrol, slurred speech, amnesia) suggest th e possibility that another medical conditi on or a substance may be causing the pan ic attack symptoms.other mental disorder s with panic attacks as an associated feat ure (e.g., other anxiety disorders and psy chotic disorders).panic attacks that occur as a symptom of other anxiety disorders a re expected (e.g., triggered by social situ ations in social anxiety disorder, by phob ic objects or situations in specific phobia or agoraphobia, by worry in generalized anxiety disorder, by separation from hom e or attachment figures in separation anxi ety disorder) and thus would not meet cri teria for panic disorder.	3
3	if an individual's presentatio n meets criteria for panic dis order and agoraphobia, both diagnoses should be assigne d.diagnostic featuresthe esse ntial feature of agoraphobia i s marked, or intense, fear or anxiety triggered by the real or anticipated exposure to a wide range of situations (crit erion a).	if an individual's presentatio n meets criteria for panic dis order and agoraphobia, both diagnoses should be assigne d.diagnostic featuresthe esse ntial feature of agoraphobia i s marked, or intense, fear or anxiety triggered by the real or anticipated exposure to a wide range of situations (crit erion a).	if an individual's presentation meets crite ria for panic disorder and agoraphobia, b oth diagnoses should be assigned.diagno stic featuresthe essential feature of agora phobia is marked, or intense, fear or anxi ety triggered by the real or anticipated ex posure to a wide range of situations (crite rion a).the diagnosis requires endorseme nt of symptoms occurring in at least two of the following five situations:1) using p ublic transportation, such as automobiles , buses, trains, ships, or planes; 2) being i n open spaces, such as parking lots, mark etplaces, or bridges; 3) being in enclosed spaces, such as shops, theaters, or cinem as; 4) standing in line or being in a crowd ; or 5) being outside of the home alone.	1, 2
4	comorbidity with depression is high also in older adults	comorbidity with depression is high also in older adults.	comorbidity with depression is high also in older adults.substances may be used as self-medication for social fears, but the s ymptoms of substance intoxication or wi thdrawal, such as trembling, may also be a source of (further) social fear.	3

5	depressive and bipolar disorders are also comorbid with separation anxiety disorder in adults.selective mutismdiagnostic criteria 313.23 (f94.0)a. consistent failure to speak in specific social situations in which there is an expectation for speaking (e.g., at school) despite speaking in other situations.b.	depressive and bipolar disorders are also comorbid with separation anxiety disorder in adults.selective mutismdiagnostic criteria 313.23 (f94.0)a. consistent failure to speak in specific social situations in which there is an expectation for speaking (e.g., at school) despite speaking in other situations.b.	depressive and bipolar disorders are also comorbid with separation anxiety disorder in adults.selective mutismdiagnostic criteria 313.23 (f94.0)a. consistent failure to speak in specific social situations in which there is an expectation for speaking (e.g., at school) despite speaking in other situations.b.the disturbance interferes with educational or occupational achievement or with social communication.c.	1, 2

The interpretations of the best model from Table 2 above are further explained here for each corresponding question number:

1. Additional information is given with descriptions of prevalence decreasing with age. The first 2 model responses give no information to infer the prevalence of separation anxiety among adults.
2. Model 1 response is not appropriate for the features of panic attacks. The Model 2 response was especially concerning as the word ‘feature’ does not appear at all in the response. Model 3 had more contextual information, though was still not perfect as it is in the Development and Course subheading rather than Diagnostic Features.
3. Models 1 and 2 contained the essential feature of agoraphobia, that is “is marked, or intense, fear or anxiety triggered by the real or anticipated exposure to a wide range of situations”. The additional contextual sentence provided in the Model 3’s response is not required here.
4. Additional information is given with the comorbidity with substance-use. The first 2 model responses give little information to infer the mechanisms of social anxiety among adults.
5. Like the interpretation of Question 3, Model 3 gives inappropriate extra information with additional criterions when a shorter answer with a simple reference to mutism as shown in Model 1 and 2 responses was sufficient.

Conclusions

Chatbots were modeled using varying NLP architectures including corpus selection (whole DSM-5 vs specific chapters on Anxiety), preprocessing with and without stop word removal as well as the inclusion of an additional contextual sentence in the output from the model. Results show that using the specific chapter on anxiety due to the much longer response time with tf-idf scoring and cosine similarity of sentences. Preprocessing with stop word removal and including an additional sentence for context in the output was also found to be most favorable based on the results. It must be noted that shorter responses may be more appropriate for some questions with the additional context not always necessary. However, a key drawback with the trained chatbot is the lack of semantic understanding of the headings as being relevant information, particularly seen with Question 2 on panic attacks. More complex modeling with headers tagged as entities, embedding representations or graph databases could alleviate this concern. Embedding representations of the corpus instead of tf- idf could capture deeper semantic information and has been shown to work well in applications for classification of medical conditions in Chinese (Zhu et al., 2019). More complex models for medical applications using graph databases are also currently in development for the corpus (Tjokro, 2017) and the queries (Ni et al., 2022).

<u>Supplementary files</u>

- fexp1.html
- fexp2.html
- fexp3.html
- anxietydsm.txt
- appi.books.9780890425596.txt

References

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders. *Diagnostic and Statistical Manual of Mental Disorders*. https://doi.org/10.1176/APPI.BOOKS.9780890425596

Bodenreider, O. (2004). The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Research*, *32*(Database issue), D267. https://doi.org/10.1093/NAR/GKH061

Center for Behavioral Health Statistics, S. (2020). *Tables 8.33 to 8.35 PE - 2020 National Survey on Drug Use and Health: Detailed Tables*.

Das, A., Sen, V., & Rose, A. C. (2022). Developing a chatbot/intelligent system for neurological diagnosis and management. *Augmenting Neurological Disorder Prediction and Rehabilitation Using Artificial Intelligence*, 273–291. https://doi.org/10.1016/B978-0-323-90037-9.00010-2

Dass, R. (2018). *Create your chatbot using Python NLTK*. Medium. https://medium.com/@ritidass29/create-your-chatbot-using-python-nltk-88809fa621d1

Ireland, D., Atay, C., Liddle, J., Bradford, D., Lee, H., Rushin, O., Mullins, T., Angus, D., Wiles, J., McBride, S., & Vogel, A. (2016). Hello Harlie: Enabling Speech Monitoring Through Chat-Bot Conversations. *Studies in Health Technology and Informatics*, *227*, 55–60. https://doi.org/10.3233/978-1-61499-666-8-55

Kaur, A., Singh, S., Chandan, J. S., Robbins, T., & Patel, V. (2021). Qualitative exploration of digital chatbot use in medical education: A pilot study. *Digital Health*, *7*. https://doi.org/10.1177/20552076211038151

Kazi, H., Pakistan Chowdhry, H. B., & Zeesha Memon, P. (2012). MedChatBot: An UMLS based Chatbot for Medical Students. *International Journal of Computer Applications*, *55*(17), 975–8887.

Lee, H., Kang, J., & Yeo, J. (2021). Medical Specialty Recommendations by an Artificial Intelligence Chatbot on a Smartphone: Development and Deployment. *J Med Internet Res 2021;23(5):E27460 Https://Www.Jmir.Org/2021/5/E27460*, *23*(5), e27460. https://doi.org/10.2196/27460

Li, A. (2019). *Mindly: A Chatbot for Helping Those with Alzheimer's Remember Loved Ones*. Medium. https://medium.com/hackmentalhealth/google-assistant-chatbot-help-with-alzheimers-12b1667334df

Ni, P., Okhrati, R., Guan, S., & Chang, V. (2022). Knowledge Graph and Deep Learning-based Text-to-GQL Model for Intelligent Medical Consultation Chatbot. *Information Systems Frontiers*, *1*, 1–20. https://doi.org/10.1007/S10796-022-10295-0/FIGURES/13

Nilsson, N. J. (2009). *The Quest for Artificial Intelligence*. http://www.cambridge.org/us/0521122937

Tjokro, M. (2017). *Developing a Medical Chatbot with Natural Language Processing and Graph Algorithm | by Moorissa Tjokro | Chatbots Life*. Medium. https://chatbotslife.com/can-a-chat-a-day-keep-the-doctor-away-f7539fea289a

Weizenbaum, J. (1966). ELIZAa computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45. https://doi.org/10.1145/365153.365168

Zhu, J., Ni, P., Li, Y., Peng, J., Dai, Z., Li, G., & Bai, X. (2019). An Word2vec based on Chinese Medical Knowledge. *Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019*, 6263–6265. https://doi.org/10.1109/BIGDATA47090.2019.9005510