

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

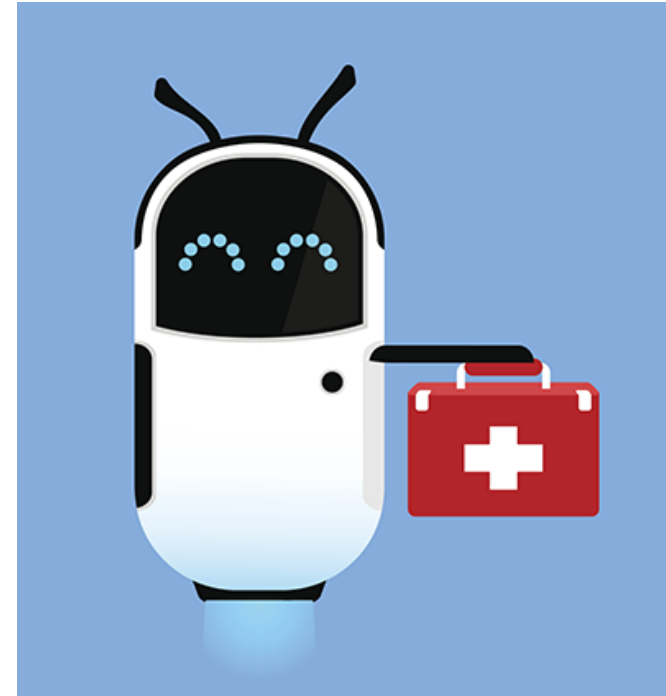
---

Apoorv Saraogee

MSDS 453 Final Project

# Introduction – business context

- Cost of treatment top concern in mental health (Center for Behavioral Health Statistics, 2020)
- Chatbots can reduce healthcare cost burden
- No known chatbots specific to medical providers in mental health
- One traditional chatbot previously done for general medical questions based on UMLS or Unified Medical Language System (Kazi et al., 2012)

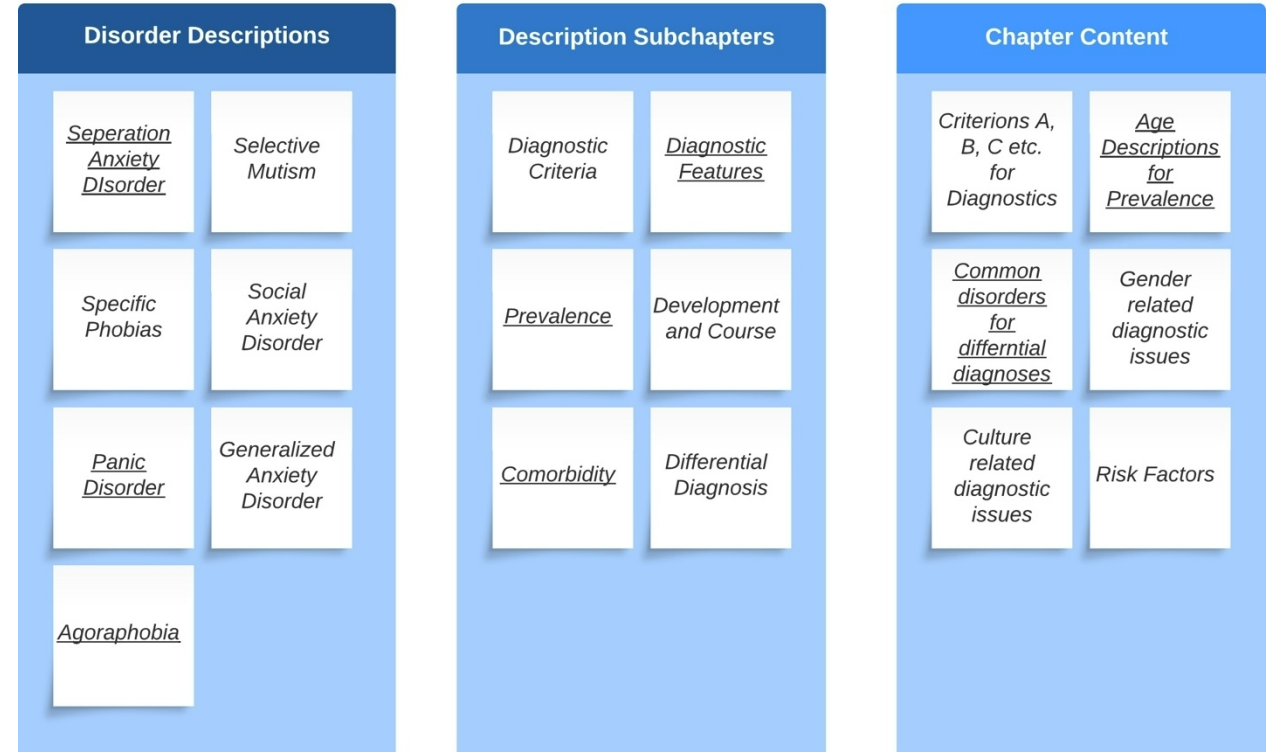


Source: Thinkstock

# Problem definition

- Lexicon/Corpus: Diagnostic and Statistical Manual for Mental Disorders, 5<sup>th</sup> Edition (DSM-5)
- Ontology for DSM-5 included in UMLS (Bodenreider, 2004)
  - Disorder Descriptions
  - Disorder Subchapters
  - Chapter Content

DSM-5 Chapter Structure with Anxiety Disorders (page 189)



# Methods

- Python kernel in Jupyter notebooks based on method from Dass 2018
  - nltk for tokenization and WordNet lemmatization
  - sklearn for tf-idf vectorization
- Cosine similarity of corpus and user input used to rank responses
- Three different NLP architectures

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

# Results

What is the prevalence of separation anxiety disorder in the United States?	What are the diagnostic features of panic attacks?	What is the essential feature of agoraphobia?	Does depression and anxiety have high comorbidity?	Is there a relationship between speaking and anxiety?
in adolescents in the united states, the 12-month prevalence is 1.6%.	panic attacks are uncommon, and unexpected panic attacks are rare, in preadolescent children.	if an individual's presentation meets criteria for panic disorder and agoraphobia, both diagnoses should be assigned. diagnostic features the essential feature of agoraphobia is marked, or intense, fear or anxiety triggered by the real or anticipated exposure to a wide range of situations (criterion a).	comorbidity with depression is high also in older adults.	depressive and bipolar disorders are also comorbid with separation anxiety disorder in adults. selective mutism diagnostic 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.

- Simplest Model 2 responses where stop words are not removed
- Somewhat relevant responses but lack of semantic understanding

# Conclusion

- Application for simple chatbot for quick lookup
- Saw better results with additional context sentence
- Lack of semantic understanding of headings in user input
  - Could be improved with embedding representations of the corpus and shown to work well for classification problems in Chinese (Zhu et al., 2019)
- Graph databases could also work well but would require domain knowledge for fine-tuning
  - Example of a UMLS chatbot where they use a Symptom Disease Matching algorithm to make a graph database (Tjokro, 2017)

# 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 (2020). *Tables 8.33 to 8.35 PE - 2020 National Survey on Drug Use and Health: Detailed Tables*.

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

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.

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>

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>