Project 3: Subreddit Classification

Sleep Apnea / Sleep Paralysis

Zhu Ye (Juliana)

Background Introduction

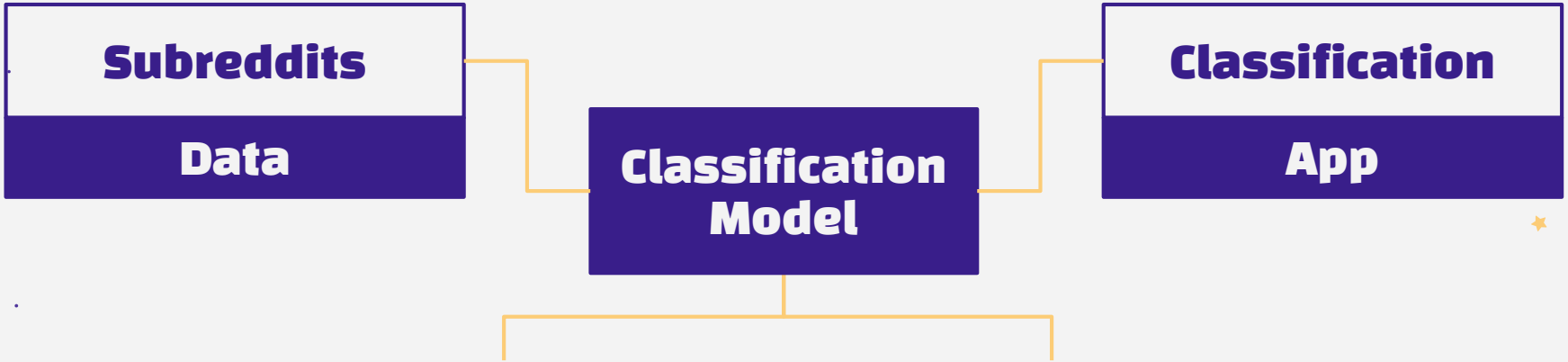
World Sleep Day 2023 March 17
Hosted by World Sleep Society
Sleep is Essential for Health

· A world annual event celebrated on the Friday before the March Equinox each year

A non-profit organization comprised of volunteers aims to educate the public about the importance of healthy sleep

Singapore Sleep Society

Problem Statement



Key words & phrases 1

What carry the most weight in our model?

Prediction 2

How is the model performance and explainability?

Sleep Apnea

Breathing interruptions during sleep:

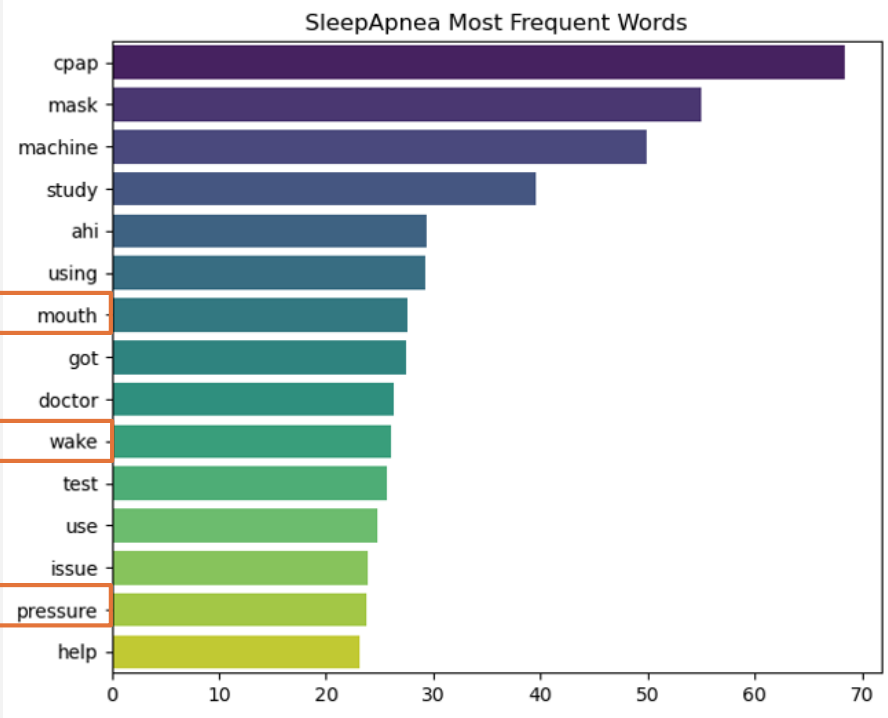
- Awakening with a dry mouth
- Gasping for air during sleep
- Morning headache
- ...

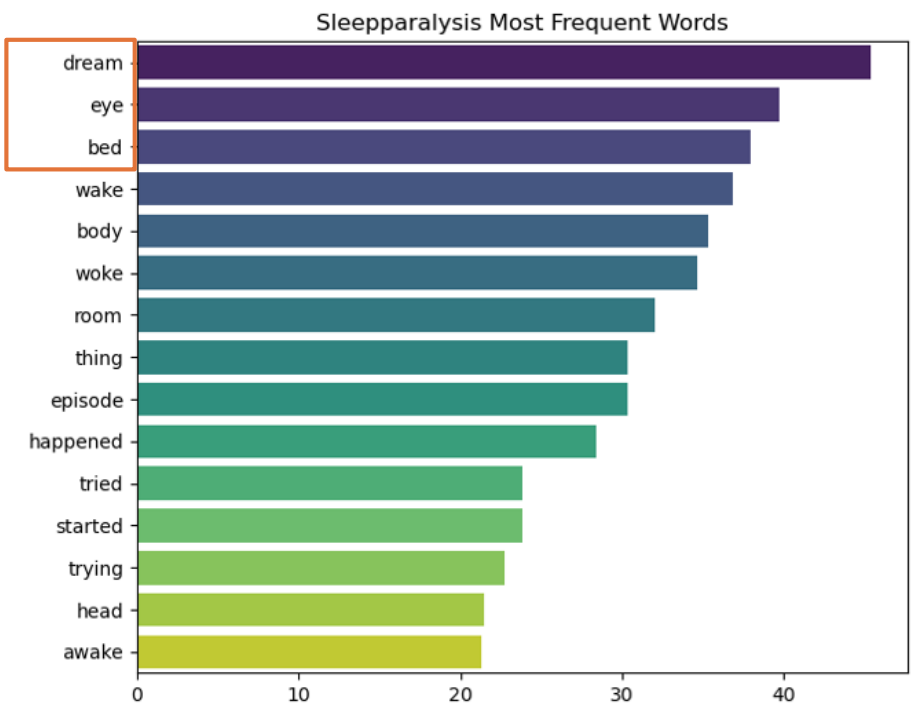
Sleep Paralysis

Temporary paralysis upon waking up:

- Awake but cannot move, speak or open your eyes
- Like someone is in your room
- Like something is pushing you down
- ...

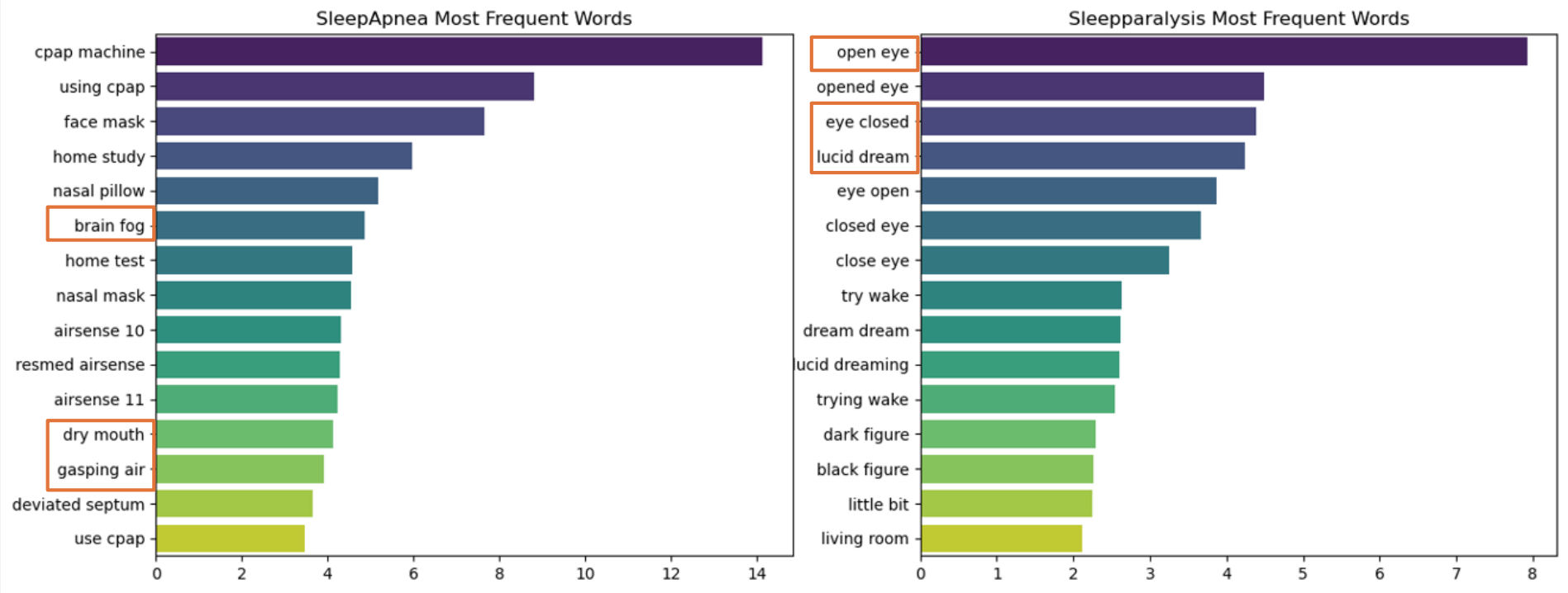
Word Frequency Analyses (Single Words)





*Plot generated by TfidfVectorizer (1-gram)

Word Frequency Analyses (Phrases)



*Plot generated by TfidfVectorizer (2-gram)

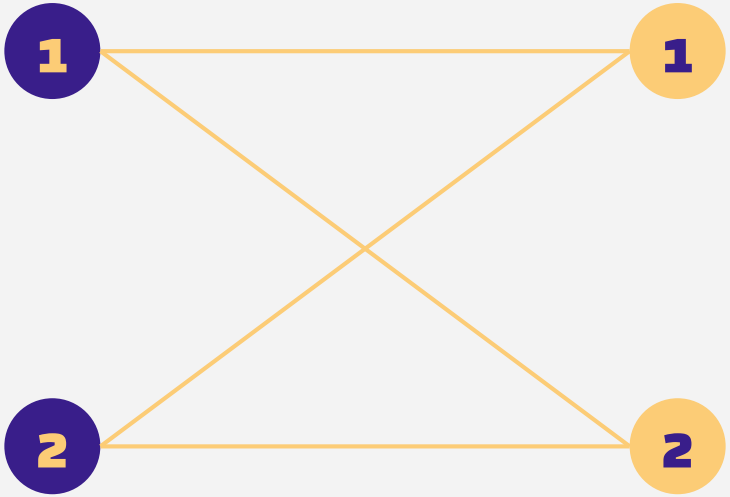
Modelling

CountVectorizer

Converts a collection of text documents to a matrix of token counts

TfidfVectorizer

Takes into account not only the frequency of words in the document but also the inverse document frequency



Naive Bayes

A probabilistic model that predicts classifications based on the probability of a document by given words

Logistic Regression

Using a logistic function to model the probability of a document belonging to a certain class

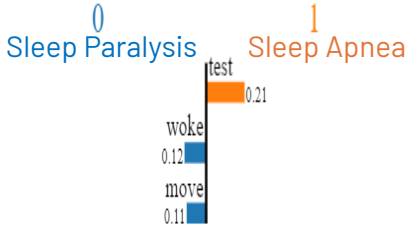
Summary of Models Performance

	Count + Naive Bayes	Tfidf + Naive Bayes	Count + Logistic Regression	Tfidf + Logistic Regression
Accuracy	0.9688	0.9666	0.961	0.9699
Precision (label: 1)	0.9675	0.9673	0.9651	0.9675
Recall (label: 1)	0.9754	0.9713	0.9631	0.9775
F1-Score (label: 1)	0.9714	0.9693	0.9641	0.9725
AUC	0.9949	0.9949	0.9937	0.9946

Model Explainability

Explain by LIME

True: 1 --> Pred: 0 | Prob: 0.71



You're basically hooked up with these wires, and if you move too much they come off, and then someone has to come in and re-hook you up. You're laying on a bed in the middle of a room with cameras staring straight at you. There's a breathing apparatus, depending on how it goes. They monitor everything, your brainwaves, your breathing, your pee (yes there's a pee test). I swear to god.

And my first thought before the test, I was like "Cool... I have nothing to hide. I wanna know if I have this. Let's see." Right?

Until I passed out... and the first time I woke up I realized I'm laying there flat on my back, in the spotlight, unable to move, on camera, all test proctor eyes on me, and I have a MASSIVE fucking boner I got in the middle of sleep sticking straight up poking obviously out of the sheets. I'm 8 inches btw. And not like in that "I'm 8 inches" kind of way that every 6 inch dude uses to try to impress his GF. I have proof.

- If we remove '**woke**' and '**move**', we can expect the model to predict "Sleep Paralysis" with a lower probability of **0.48** (0.71 - 0.12 - 0.11), i.e. to predict "Sleep Apnea" correctly with a probability of **0.52** (1-0.48)

Explain by Logistic Regression

Word	Coefficient	Odds Ratio
test	2.165	8.715
move	-3.268	0.038
woke	-1.738	0.176

- An increase of one unit for the word '**test**' is associated with a **8.72**-fold increase in the odds of 'Sleep Apnea'
- An increase of one unit for the words '**move**' and '**woke**' is associated with a **26.25**-fold (1/0.038) and **5.68**-fold (1/0.176) increase in the odds of 'Sleep Paralysis', respectively

**This is while holding all other variables constant in the model.*

Conclusion and Recommendation

We created a web application that provides immediate feedback to users about potential sleep disorders according to their reported sleep experiences.

Key words & phrases 1	Prediction 2
Keywords *Sleep Apnea*: mouth, wake and pressure *Sleep Paralysis*: dream, eye and bed **Phrases** *Sleep Apnea*: brain fog, dry mouth and gasping air *Sleep Paralysis*: open eye, eye closed and lucid dream	We developed a reliable classification model with an accuracy of 0.97 that predicts the likelihood of individuals having Sleep Apnea / Sleep Paralysis based on users' experiences shared on Reddit and the predicted result is explainable.

Demo Time