

Topic modeling for depression drug reviews

Model's used : Gensim and NMF

Results : The **Gensim** model kept giving unclear and overlapping topics even after further data preprocessing such as stemming and getting rid of frequent but less meaningful words by adding them to stop word list.

The **NMF** model gives the following results , which show a better pattern between topic and terms. From this result I can conclude the following topics

Topic 1 = When patients start taking the medication

Topic 2 = Symptoms

Topic 3 = Drug names

```
display_topics(nmf_model, cv.get_feature_names(), 13, topic_names=["topic_1","topic_2","topic_3","topic_4"])
```

```
Topic: ' topic_1 '  
talking, startled, day, effectsa, feminine, weeksnot, mgof, days, monthsin, medicationjust, weekly, nightmakes, sleepier
```

```
Topic: ' topic_2 '  
depression, anxiety, effectsa, helpful, helpjust, medicationjust, sew, monthsin, panicky, mgof, prescriptions, weeksnot, attack  
s
```

```
Topic: ' topic_3 '  
feelingquot, better, feminine, weeksnot, feelings, goodjust, live, weekly, thinkingcompulsive, normalizes, day, happynonsuicida  
l, depressed
```

```
Topic: ' topic_4 '  
yearseasonally, lifefor, timei, triggered, ago, effexorvenlafaxine, doctoral, wellbutrinuntil, weightloss, prozacseems, zolpide  
m, effectsa, workers
```

* Further data preprocessing will be done to improve the generalization of the topics