

Paper DH13

Exploring Antidepressant Side Effects in the UK Through Topic Modelling and Social Media Analysis

Shreya Mandot, Amna Dridi, Zebedee Wanyonyi*, Edlira Vakaj – **Birmingham City University**

ICC Birmingham – November 8th, 2023



Agenda

- Overview
- Aim and Objectives
- Research Questions
- Existing Work
- Proposed Methodology
- Dataset
- Steps for Proposed Study
- Results
- Conclusion & Future Work



Research Problem

- **‘Antidepressants’**, known to have inconsistent effects on different people with different medical situations, making it difficult to determine how well they work to treat any mental illness.
- Current trials to measure the **side effects** of antidepressants are time consuming and pretty biased!
- **Social media platforms** let people to communicate with others, their daily problems. As a result, it can serve as a discrete, large-scale, and naturalistic data source for researchers studying these side effects of antidepressant on users.

Aim & Objectives

Aim

To understand the side effects of antidepressants from the social media perspective, using topic modelling and Natural Language Processing

Objectives

- Conduct a thorough review of the literature on antidepressant side effects research in the psychiatry domain.
- Data Collection for antidepressant usage effectiveness.
- Extracting important features from Tweets using NLP techniques.
- Automated extraction/ prediction of side effects of antidepressants using Topic modelling.
- Evaluate the results and outcome of topic modelling techniques and create a benchmark of antidepressants' side-effect.
- Compare the variation in results for the different number of topics with the help of coherence score and perplexity score.

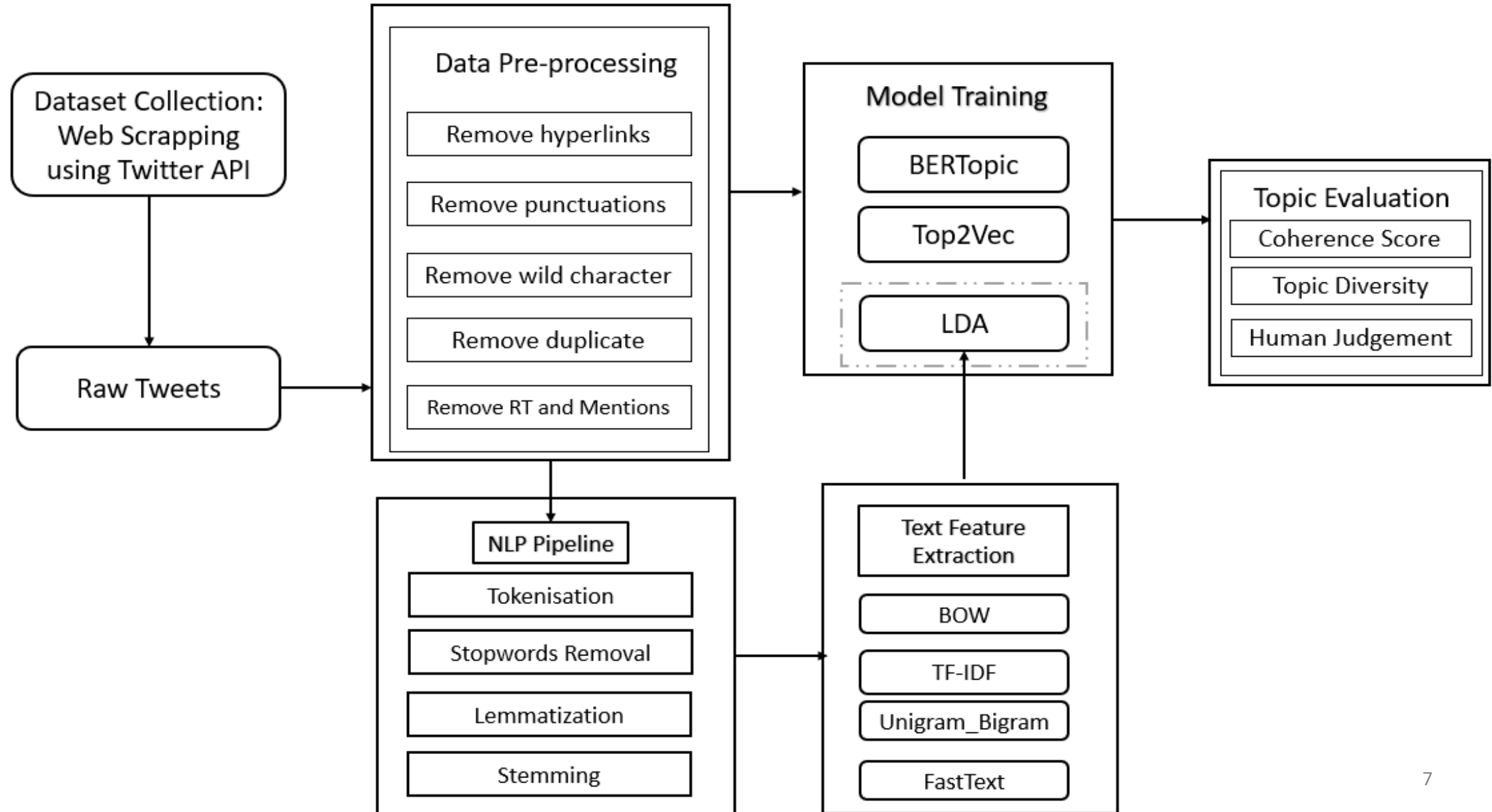
/ Questions Asked...

- ✓ How would Natural Language Processing (NLP) techniques be able to represent side effects of drugs in the psychiatry domain from social media platforms?
- ✓ How do Machine Learning (ML) and Topic Modelling techniques perform in accurately detecting the side effects of antidepressants?
- ✓ Can we trust the insights we get from the social media platform?

Existing Work

Research Paper	Dataset	Techniques Applied
Beyond LDA: Exploring Supervised Topic Modelling for Depression-Related Language in Twitter	3 Million Tweets	LDA based of n grams, SLDA with bag of word, Supervised Anchor Model(on LDA)
UNSUPERVISED DYNAMIC TOPIC MODEL FOR EXTRACTING ADVERSE DRUG REACTION FROM HEALTH FORUMS	Ask a patient dataset	HAN, FastText, word2Vec, bag of n gram-NMF
A Topic Modelling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts	50k tweets	TF-IDF. LDA, NMF, Top2Vec, Bertopic
Understanding Side Effects of Antidepressants: Large-scale Longitudinal Study on Social Media Data	Twitter Data- Used 297 names of medicine from FDA	Pre-treatment data: unigram feature selection and logistic regression. Post-treatment data- Sparse Additive Generative (SAGE) model with n-grams
Depression detection from social network data using machine learning techniques	Facebook data	Decision Tree, SVM and KNN (Labelled dataset)
USE OF TWO TOPIC MODELING METHODS TO INVESTIGATE COVID VACCINE HESITANCY	Tweets	LDA, Top2Vec
An Effective BERT-Based Pipeline for Twitter Sentiment Analysis: A Case Study in Italian	Tweets	BERT for Twitter sentiment analysis
Top2Vec: Distributed Representations of Topics	20 News Groups Dataset, Yahoo Answers Dataset	LDA and PLSA using bag-of-words , Top2Vec
Consumer perceptions of telehealth for mental health or substance abuse: a Twitter-based topic modelling analysis	Tweets from January 2014 to June 2021	Comparison of pre pandemic and post pandemic tweets using BERT

Workflow



Methodology

- **Latent Dirichlet Allocation (LDA)**
 - Bag of Words
 - Term Frequency – Inverse Document Frequency
 - Unigram-Bigram
 - FastText
- **BERTopic**
 - C TF-IDF, BERT
- **Top2Vec**
 - Universal-Sentence-Encoder Word Embedding



Dataset

- Food and Drug Administrator Approved list of Antidepressant Drugs.
- Used Tweeter Application Processing Interface to retrieve public Tweets.
- Queried 9500 Tweets in English language.

I caught cold, having toothache and headache, on my period, and not to mention that I actually might have PMDD cuz I have the history as Escitalopram consumer *crying snuggled up in blanket on my bed

starting lexapro soonðŸ˜˜ðŸ˜˜(Donâ€™t tell me ur bad experience with it because i have severe hypochondria and will shut and cry and puke)â–ĩ,â–ĩ,

@RunInWithATree I take Lexapro and being without it really messes you up. I felt a lot of dizziness when I was getting used to it. Are you out of refills? Might be making the sleep deprivation worse.

@Poly40399120 I miss taking Sertraline, since they moved me onto Escitalopram my libido has been almost non-existent and I have a hard time trying to stay focused...

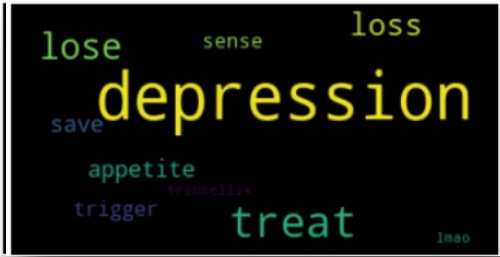
My mirtazapine dosage has been 15mg for almost 4 weeks and my sleep is still awful so I'll hopefully be having a GP telephone appointment today to see if I can lower it anymore (I'm hoping I can go down to 7.5mg but idek if that's possible)

agomelatine, amineptine, amitriptyline, amoxapine, bupropion, butriptyline, citalopram, clomipramine, desipramine, desvenlafaxine, dibenzepin, dosulepin, doxepin, duloxetine, escitalopram, etoperidone, fluoxetine, fluvoxamine, hydroxynefazodone, imipramine, iprindole, levomilnacipran, lofepramine, maprotiline, mazindol, meta-chlorophenylpiperazine, mianserin, mirtazapine, nefazodone, nisooxetine, nomifensine, norclomipramine, northiaden, nortriptyline, opipramol, oxaprotiline, paroxetine, protriptyline, reboxetine, sertraline, setiptiline, trazodone, triazoledione, trimipramine, venlafaxine, vilazodone, viloxazine, vortioxetine, zimelidine

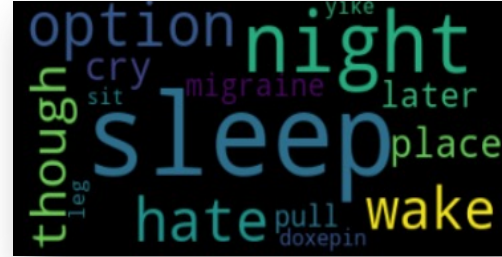
Topic Modelling Results

Topic Modelling	Feature Extraction	Coherence Score with 35 topics	Coherence Score with 65 topics	Topics Created on Own
LDA	BOW	0.47	0.52	NA
LDA	Unigram Bigram	0.48	0.56	NA
LDA	TF-IDF	0.49	0.51	NA
LDA	FastText	0.60	0.74	NA
BERTopic	BERT, TF-IDF	0.56	0.63	0.71
Top2Vec	Universal-Sentence Encoder	NA	NA	0.47

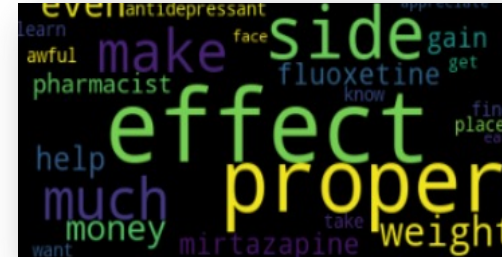
Visualizations



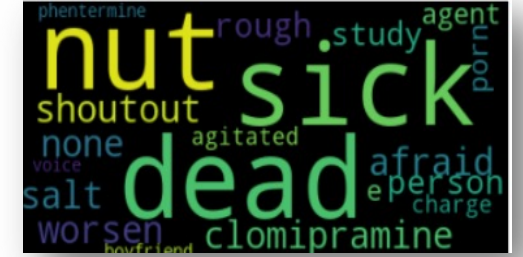
LDA-BOW



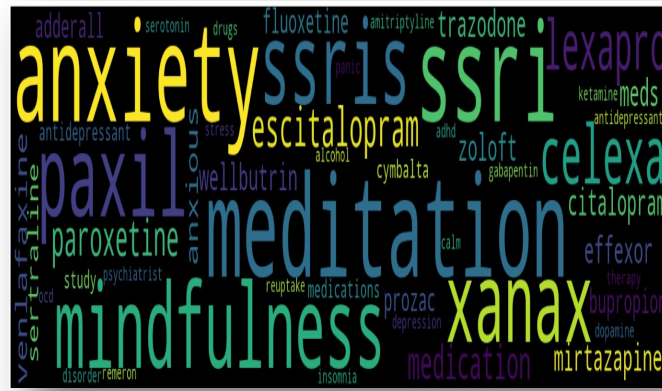
LDA-Uni-Bigram



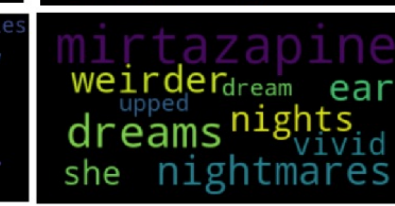
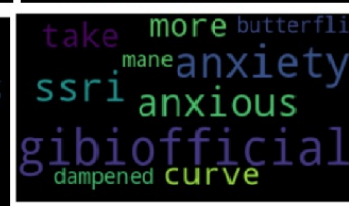
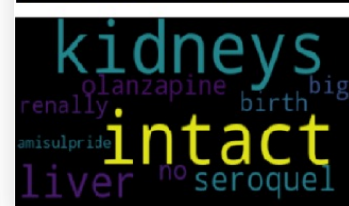
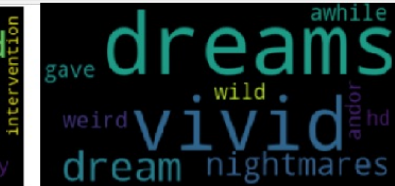
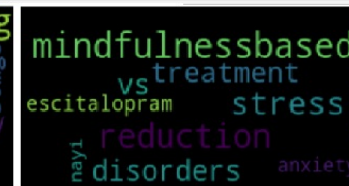
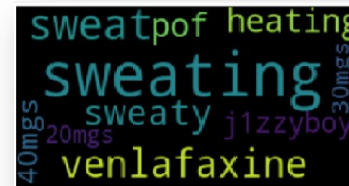
LDA-TF-IDF



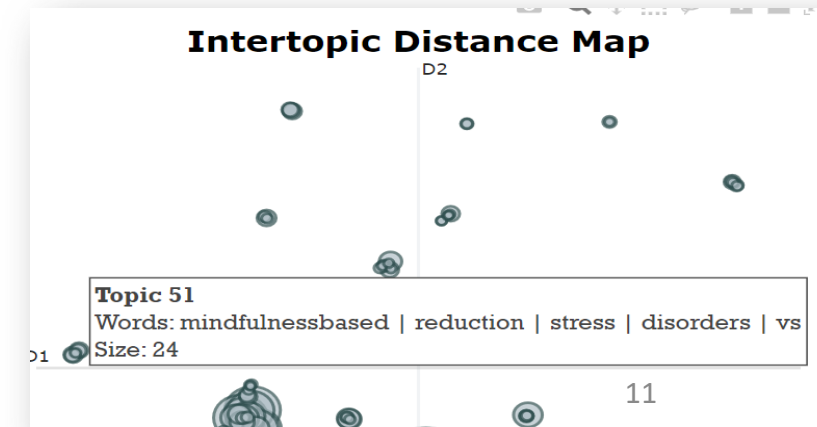
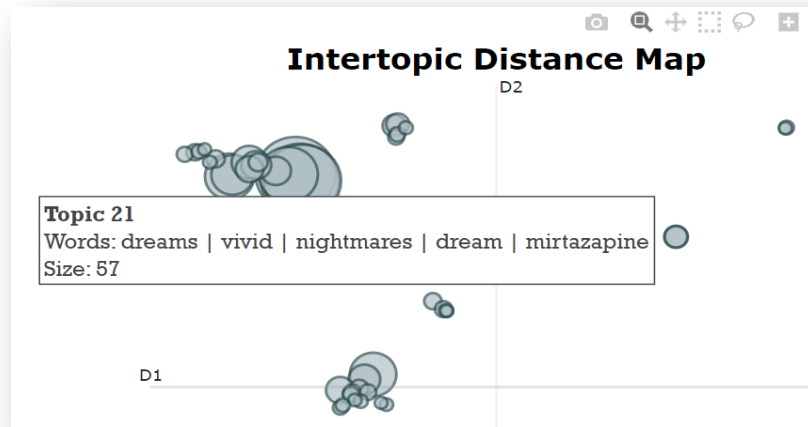
LDA-FastText



Top2Vec



BERTopic

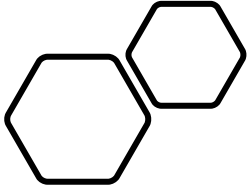


Conclusions

- Amongst all Topic Modelling techniques used - BERTopic proves to be best
- Side effects predicted are amongst the ones appearing on 'FDA Approved List'
- Key Side-effects extracted : Weight gain, Nausea, Migraine, Anxiety, Depression, Sweating, Weird dreams, etc.
- Getting excellent Coherence Scores not necessarily implies having an accurate representation of distributional information concerning human evaluations.

Future work

- Extend Dataset
- Extend Research Analysis to,
 - Other language tweets (apart from English)
 - Other Social Media Platforms (with same topic modelling)
- Implement new topic modelling techniques



Thank you

Questions?

Contact Details:

- Amna Dridi: amna.Dridi@bcu.ac.uk
- Shreya Mandot: shreya.mandot@mail.bcu.ac.uk