Using deep learning methods for discovering associations between drugs and side effects based on topic modeling in social network

Article *in* Social Network Analysis and Mining · December 2020

DOI: 10.1007/s13278-020-00645-8

Behnaz Eslami
Islamic Azad University Tehran Science and Research Branch

Zahra Rezaei
University of Kashan

Majid Fouladian
Islamic Azad University - Saveh Branch

Social Network Analysis and Mining (2020) 10:33
https://doi.org/10.1007/s13278-020-00645-8

ORIGINAL ARTICLE

Using deep learning methods for discovering associations between drugs and side effects based on topic modeling in social network

Behnaz Eslami[1] · **Zahra Rezaei**[2] · **Mehdi Habibzadeh**[3] · **Majid Fouladian**[4] · **Hossein Ebrahimpour-Komleh**[2]

Received: 9 November 2018 / Revised: 9 November 2018 / Accepted: 6 May 2020

Abstract

The relationship between drug and its side effects has been delineated in two websites, namely Sider and WebMD. The aim of the present paper is to find the relationship between drug and its side effects as reported by typical users of a website called *Ask a patient*, and to compare these reports with the side effects in reference sites. In addition, the typical users' comments on highly-commented drugs (neurotic drugs, anti-pregnancy drugs and digestion drugs) within last decade were analyzed. The reason for such investigation is the fact that typical users' comments and their tendencies can be considered as an important factor in determining the best drugs in improving them or decreasing their risk dangerous. Typical users' comments on drugs' side effects were gathered from the website *Ask a patient.* Then, the data on drugs (neurotic drugs, anti-pregnancy drugs and digestion drugs) were classified according to deep learning model. At first using the model, the three issues, namely drug, its side effect and the cause of the side effect, were explained. Afterward, using topic modeling, the main topics of side effects for each group of drugs were identified. Finally, using the websites of Sider and WebMD in which the side effects of drugs are reported, the side effects of the three classes of drugs were retrieved. The goal of the present research was to analyze typical users' comments reported on the website called *Ask a patient*, and to compare these comments with the reports about the side effects of drugs from important sites. Our model demonstrates its ability to accurately describe and label side effects in a temporal text corpus. By taking full advantage of deep learning classifiers, the used methods in text mining is shown to be accurate and effective for discovering association between drugs and side effects. Moreover, through combining with modular classifier in addition to topic modeling, this model has the capability to immediately locate information in reference sites to recognize the side effect of new drugs. In fact, due to the unbiased nature of typical users' comments, these comments can be a reliable indicator for drug producer companies to reduce the side effects of drugs.

Keywords Deep learning · NLP · Classification · Topic modeling · Text mining

✉ Behnaz Eslami
behnazeslami30@gmail.com

Zahra Rezaei
z.rezaei2010@gmail.com

Mehdi Habibzadeh
me_habi@encs.concordia.ca

Majid Fouladian
fouladi@iau-saveh.ac.ir

Hossein Ebrahimpour-Komleh
ebrahimpour@kashanu.ac.ir

1 Department of Computer Engineering, Science and Research Branch Islamic Azad University, Tehran, Iran

2 Department of Computer and Electrical Engineering, University of Kashan, Kashan, Iran

3 Data Science Department, P/S/L Consulting Group, Montréal, Québec, Canada

4 Department of Electrical Engineering, College of Technical and Engineering, Saveh Branch, Islamic Azad University, Saveh, Iran

1 Introduction

The unwelcome, undesirable or dangerous effects that a drug may have are referred to as adverse drug reaction (ADR, or adverse drug effect). The term 'side effect' which is inexact term refers to inadvertent, secondary effect that is observed along with the therapeutic function of the drug. Side effect may fluctuate from person to person.

Adverse drug reactions can be considered a form of toxicity, i.e., due to accidental or intentional overdose, or other causes of elevated blood levels, and drug interactions. All drugs have the potential for adverse reactions; it is assumed that between 3 and 7% of all hospitalizations are due to adverse drug reactions.

A systematic review 25 prospective observational studies demonstrated that 5.3% of all patients are faced with adverse drug reactions (Kongkaew et al. 2008). Thus, early detection of these events could greatly impact on human health. According to the Agency for Healthcare Research and Quality report, annually over 770,000 of people are injured and/or die in hospitals due to adverse drug reactions (Classen et al. 1997). Thus, societies required alternative approach to detect side effects of the clinical medications. Economically, ADRs can considerably increase patient's hospitalities costs (Bordet et al. 2001; Sultana et al. 2013).

The data from social media are a novel and rich source of data using which the trends and thinking flow of users about side effects of drugs and special events in the field of health can be identified and managed. The purpose is to use these data to help patients.

Ask a patient (2001) is the website that empowers patients by allowing them to share and compare medication experiences and was awarded the 2012 Webby Award for best website in the Pharmaceutical Category. The *Ask a patient* database contains more than 4000 chemically prepared prescription drugs approved by FDA's Center for Drug Evaluation and Research. You can find comment of prescription or over-the-counter drugs, based on fine-tuned search criteria (age, gender, symptom, etc.). The difference between written and oral language in social media can create noise.

In addition, lack of a suitable structure and imbalance data in drug groups are considered as important challenges in classifying data from social media. So, in spite of richness in health-related data in social media, little practical use of these data is made. The method used in the present research had three main phases: first, in order to extract features from social media, a learning process happens automatically in deep learning. The comments by the users of a website *Ask a patient* were processed to describe side effects and accordingly to reduce the difference between written and oral language as well as the noise.

Second, the efficiency of deep learning method in classifying the data from the website *Ask a patient* was proved. The results showed that deep learning performance accompanies high accuracy and high speed. Third, the advantages and disadvantages of using comments in recognizing side effects were compared with the side effects reported in the comments in two websites of Sider and WebMD. In so doing, deep learning method HAN (Yang et al. 2016) was employed to classify users' comments. Then, for determining specific topic in each group of drugs, the non-monitoring method (NMF) of topic modeling was employed.

2 Related work

Sarker and Gonzalez (2015) highlighted importance of employing advanced NLP-based information generation in accuracy of ADR sentence detection and classification by used traditional text classification such as Support Vector Machine, Naïve Bayes and Maximum Entropy.

Ginn et al. (2014), they presented an annotated Twitter corpus focused on ADR mentions with broad. They applied two supervised machine learning approaches (NB and SVM) on broad range of annotated medications related to ADR tweets in Twitter. Although, the classifier shows moderate performance, but it was considered as fundamental for future development of advanced techniques. In line with this approach, Akhtyamova et al. (2017a) used convolutional neural networks (CNN) model with word2vec embedding for Twitter comments classification. In contrast to Arker's model (reference), their proposed model not only uses a small fraction of features for gathering information, but show high performance of applicability in text classification.

Lee et al. (2017) suggested a semi-supervised CNN-based framework for adverse drug events (ADE) classification in Twitter. A Twitter datasets used in PSB 2016 Social Media Shared Task applied for evaluation of model, resulting high performance classification of ADE with +9.9% F1-Score. Notably, Adverse Drug Event detection (ADE) surveillance systems required small number of labeled instances.

Akhtyamova et al. (2017b) present a CNN-based architecture consisting of numerous parameters to predict revealing adverse drug reaction based on the quantity of vote. To evaluate the performance of model, a large-scale medical dataset derived from medical websites was utilized. In contrast to previous reports networks, the proposed end-to-end model does not require handcrafted features and data preprocessing, resulting an enormous improvement for standard CNN-based methods.

In this study aimed to investigate the written topic modeling of typical users and to identify the changes in reporting comments within 10 years. In a way so that the designed model can provide researchers with immediate capability of analyzing comments through combining deep learning methods. The inclination of the comments within years showed a significant change of users' comments in reporting side effects of drugs. This reduction can be attributed to using drug supplements, change in life style, genetic improvement of drugs, etc.

3 Methods

3.1 Workflow of the research

This paper is organized in two sections as Classification and Topic Modeling (Fig. 1).

4 Section 1: Classification

4.1 Data sources

Prior to collecting data, we selected a set of drugs of interest, which were likely to have a large number of associated comments in *Ask a patient* database. We selected drugs that were prescribed for chronic diseases and syndromes for which large numbers of comments were expected and drugs with high prevalence of use. The names of the medications are reported in separate classes (Anti-depressant Medicines, Anti-Pregnancy Medicines and Digestion Medicines) in Figs. 11, 12, and 13 in Appendix.

4.2 Pre-processing

The pre-processing of comments in both data is done as follows:

1. Data shuffling
2. Converting all uppercase words to lowercase
3. Elimination of special characters like: @, !,/, *, $ and etc.
4. Remove stop word: at, of, the, …
5. Correction of words with repeated characters like: pleaseeeeeeeeee and/or yessss
6. Convert contractions to base format like: I'm → I am
7. Lemmatization: I **started taking** almost two months ago. → I **start take** almost two months ago.

4.3 Cross-validation

For many classification models, the complexity may be governed by multiple parameters. In order to achieve the best prediction performance on new data, we wish to find appropriate values of the complexity parameters that lead to the optimal model for a particular application.

If data are plentiful, then a simple way for model selection is to divide the entire data 1 into three subsets, the training set, the validation set and the test set. A range of models are trained on the training set, compared and selected on the validation set, and finally evaluated on the test set.

Among the diverse complex models that have been trained, the one having the best predictive performance is selected, which is an effective model validated by the data

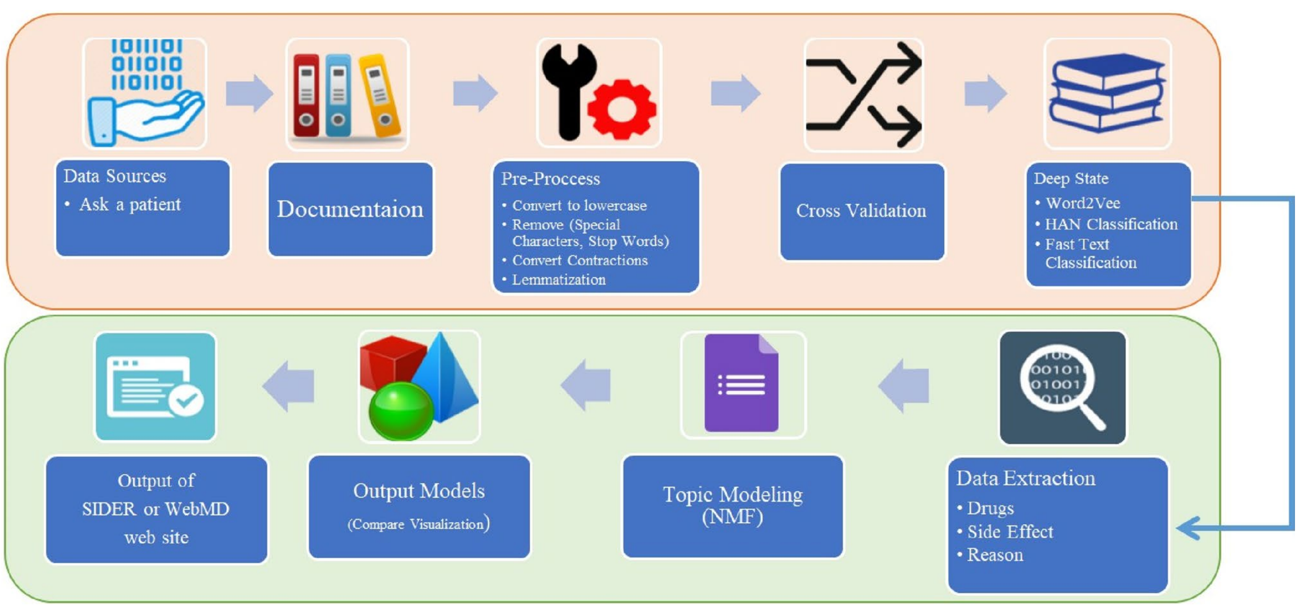


Fig. 1 The workflow of the proposed deep learning-based strategy is illustrated

in the validation set. In a practical application, however, the supply of data for training and testing is limited, leading to an increase in the generalization error. An approach to reducing the generalization error and preventing over-fitting is to use cross-validation. The distribution of data for each group is shown in Table 7 in Appendix.

4.4 Deep classification

The used methods for data classification are including HNN (Yang et al. 2016) and FastText (Joulin et al. 2017) with similar word2vec section. Once word2vec generated, this file used for further steps of study.

4.4.1 HAN method

Hierarchical Attention Network (HAN) has two distinctive characteristics: (I) a hierarchical structure that mirrors the hierarchical structure of documents; and (II) two levels of attention mechanisms applied at the word and sentence-level, enabling it to attend differentially to more and less important content when constructing the document representation. In addition to these, the HAN network composed of quite a few parts including, a word sequence encoder, a word-level attention layer, a sentence encoder and a sentence-level attention layer. HAN hypothesized that considering sentence and documents structure in modeling play positive role in better representation of document structure in the model architecture (Table 1).

4.4.2 FastText method

This method proposes a simple and efficient approach for classifying the texts and its expressions. Large number of researches shows that the rapid classification of text with this method is faster in comparison with deep learning methods in terms of accuracy and using commands for training and evaluation (Table 1). Architecturally, there are two major and influential differences:

1. **Softmax:** is a hierarchy, based on the Huffman-encoded tree structure that reduce Time Complexity **O(Kd)** to **O(d log k)** where **K** is the number of targets and **D** the dimension of the hidden layer.
2. **N-gram features:** While Bag of words is invariant to word order but taking explicitly this order into account is often computationally very expensive. Instead, we used bag of n-gram as additional features to capture some partial information about the local word order. This is very efficient in practice while achieving comparable results to methods that explicitly use the order.

4.5 Evaluation metrics

Precision (positive predictive value) and recall (sensitivity): are appropriate fraction of retrieved related samples from all and relevant instances, respectively. Application of these metrics depends on understanding and measuring of relevance.

Accuracy: The accuracy criterion is the accuracy of the x-group classification against all items where the x-tag for investigating records is suggested by means of classification. This criterion indicates how much classification output is trustable.

F-measure: This criterion is a combination of call metrics and accuracy and it is used in cases where it is impossible to consider special importance to each of the two criteria.

Kappa: This criterion is often used to test the reliability of the viewer and to compare the accuracy of the system in terms of how much generated output is coincidental (Table 2).

5 Section 2: Topic Modeling

5.1 Input datasets in the second phase

The three classes of drugs in the time period between 2008 and 2018 are used based on what is presented in Figs. 11, 12, 13.

Table 1 (HAN and FastText) training phase configuration

Training phase
Initializations:
Configuration of distributed parameters {Device: {**NVIDIA GEFORCE GTX 1050, RAM 16G**}}
Configuration of optimization {Name of optimization: {"**Adam**"}} **configuration of loss** {Name of loss-function: {"**Sigmoid**"}}
Initials {Pad_Seq_Len: {**150**},
Embedding_Dim: {**100**},//for creating Word2Vec model
Batch_Size: {**32, 64 and 128**},
Epochs: {**100**}},
Learning Rate: {**0.1, 0.01, 0.001**}
Configuration of dataset {Datasets: {**Train.json**}}
Main ():
Select the Dataset//Based of Application and select Train part
Select the Network//A function that applies the model to a batch of documents, returns last op containing the log predictions
Create a dataset provider that loads data from the dataset//Return [Content, Label]
Create training operations
Run the training

5.2 Topic modeling

As a linear-algebraic model, non-negative matrix factorization (NMF) includes high-dimensional vectors into a low-dimensional image. NMF like principal component analysis (PCA) considers the fact that the vectors are non-negative. NMF through including the vectors into lower-dimensional form causes the coefficients to be non-negative, as well.

Using the original matrix A, the two matrices of **W** and **H** can be obtained so that **A = WH**. As NMF has an inborn clustering property and **W** and **H** represent the following information:

A (Document-Word Matrix): input that shows which words appear in which documents.

W (Basis Vectors): the topics (clusters) elicited from the documents.

H (Coefficient Matrix): the membership weights for the topics in each document.

W and **H** are calculated by optimizing an objective function (like the *EM algorithm*), updating both **W** and **H** iteratively until they are converged. In this way, the NMF topic modeling configuration is provided in Table 3.

6 Results

6.1 Usage model

In this research, we used user's comments of *Ask a patient* to extract side effects of drugs. In the field of deep learning, the following issues are considered in the training phase.

Table 2 Evaluation metrics formula

$Precision = \frac{TP}{TP+FP}$
$Recall = \frac{TP}{TP+FN}$
$Accuracy = \frac{TP+TN}{TP+TN+PF+PN}$
$\text{F-}Score = \frac{Precision*Recall*2}{Precision+Recall}$
$Kappa = \frac{Pr(a)-Pr(e)}{1-Pr(e)}$

In general, the size of a window that moves on texts in both FastText and HAN methods is called *Pad_Seq_Len*, and we considered quantity equal to 150 because generally the maximum size of comments is 150 where the length of sentences and semantic conjugation are important. Moreover, the value of Embedding dim was 100 (Table 4). We evaluated several optimizations such as *Stochastic Gradient Descent*, *RMS prob* and *Adam*. That *Adam* shows better results.

Also in Section 2, for extracting critical topics *ngram_range* to detecting words in a define scale and *min_df* to finding words in documents by minimum frequency were determined.

The value of *ngram_range* choose based on the side effects expressions that extracted from Sider or WebMD website, although other values such as (1,2), (2,3) and (3,3) were determined but (2,2) was the best choice (Table 5).

Table 3 Topic modeling configuration

NMF topic modeling
Initializations:
Configuration of distributed parameters {Device: {**CPU - Core i7, RAM 16G**}}
Number of topics: {**10**}
Number of top words: {**20**}
Configuration of feature extraction by using **TfidfVectorizer**: {
Initials: {
ngram_range: {**(2, 2)**},
min_df: {**2**},
Configuration of NMF topic modeling parameters and **fit** by **TfidfVectorizer:** {
n_components: {**Number of topics**}
init: {'**nndsvd**'},//better for sparseness
}},
Run to extracting topics

Table 4 The hyper parameters in training phase

Pad_Seq_ Len	Embed-ding_dim	Dropout_ Keep_Prob	Loss	Optimization
150	100	0.5	Sigmoid	Adam

Table 5 NMF topic modeling parameters

ngram_range	min_df
(2, 2)	2

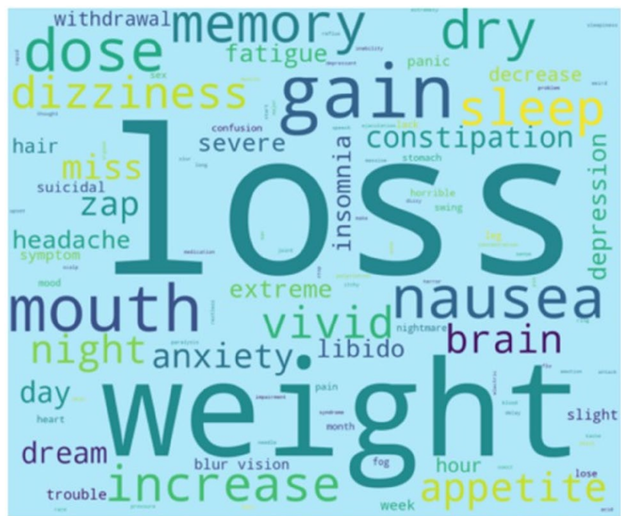


Fig. 2 Anti-depressant topic modeling visualization (anti-depressant topic modeling (*Ask a patient)* is reported in Table 10 in Appendix)

6.2 Implementation model in section 1

In this research, the used hardware includes: *NVIDIA GEFORCE GTX 1050* and *CPU Intel Core i7*. Two methods of classification were applied against three different data groups listed in the following table. As shown the best results in Table 6, in each method the learning rate and batch size were evaluated and different criteria have been tested for each type of model according to the type of data, and various values have been obtained. For example, applying HAN method included with batch size of 128 and learning rate 0.001 on *Ask a patient* dataset resulted highest accuracy (0.924). Confusion matrix HAN and FastText; for best results are reported in Tables 8 and 9 in Appendix.

6.3 Implementation model in section 2

Considering the output of the previous phase, the three features, namely side effects, reason and drug, were used. Accordingly, in each class of drugs (neurotic medicines, anti-pregnancy and digestion), 10 topics of high priority

Table 6 Best result of deep learning classification methods on dataset

Dataset	Method	Batch size	Learning rate	Accuracy	Kappa	Recall	precision	F1-score
Ask a patient	HAN	128	0.001	0.924	0.885	0.921	0.926	0.923
	FastText	128	0.001	0.909	0.863	0.908	0.909	0.909

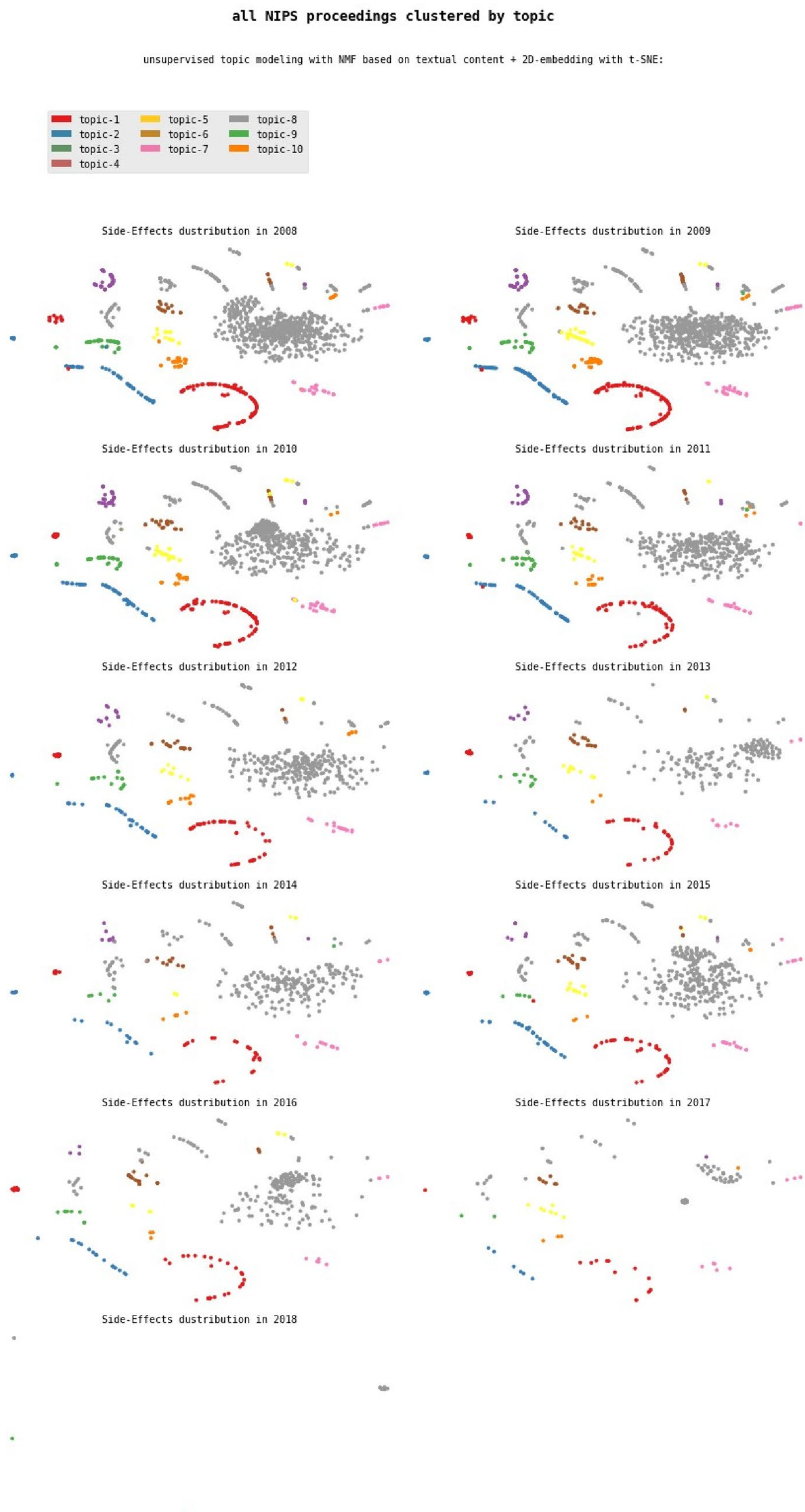


Fig. 3 Scatterplot of anti-depressant medicines topics on the website of *Ask a patient* based on year

Fig. 4 Anti-pregnancy topic modeling visualization (anti-pregnancy topic modeling (*Ask a patient*) is reported in Table 11 in Appendix)

were selected. As shown in Tables 10, 11 and 12, topics of each class are verbally similar. After extraction of these tables, all are mapped with a similar word, and meaningless topics were deleted. Figures 2, 4, and 6 show the frequency of repetition of topic models, and Figs. 3, 5 and 7 show the dispersion of topics on the website of *Ask a patient* during the years 2008 to 2018. The users' comment about side effects shows a different model in each year.

According to Figs. 8, 9, and 10, users' comments were different from the side effects of drugs reported in the websites Sider and WebMD in case of the three classes of drugs; however, the websites had reported some side effects but with a low frequency. The blue diagram shows the frequency of side effects reported in websites, and the red diagram presents the comments by typical users from side effects; however, some reports overlapped with the users' comments and the websites (Sider and WebMD) in terms of topics.

7 Discussion

In the present research, the deep learning methods of HAN and FastText were employed to classify side effects of three classes of drugs, namely neurotic, anti- pregnancy and digestion. Due to the fact that the comments on these three classes of drugs had a high frequency, they were investigated. In the first phase, the extracted data from the website *Ask a patient* were entered into the model. Then, in the pre-processing phase special characters, sign and stop words were removed and the characters were converted into small-case letters in order to improve the text. In the second phase, the three fields of drugs, the side effect and the reason of side effect were investigated. Then, these data were exposed to classifying phase (topic modeling) to extract 10 topics of high priority from the three groups of drugs. The outputs show that the frequency of occurrence of side effects reported in the comments from *Ask a patient* was different from the side effects reported from Sider and WebMD, and in some minor cases some similarities in frequency were seen. Finally, the proposed model is compared with the output of drug's side effects, and the analysis of users' reports and the sites' reports is illustrated. In addition, the users' trends about side effects were analyzed for the time period between 2008 and 2018 during 10 years. As it is clear, the users' comments have changed gradually.

Finally, the obtained results derived from the preliminary analysis of drug classification presented in confusion matrices and interpreted using accuracy rate and false positive ratio.

In this work, we used a simple method for text classification by deep learning models. In contrast to unsupervised trained word vectors derived from word2vec, our word features can be averaged together to generate appropriate sentence representations.

In comparison with recent deep learning-based methods, the FastText and HAN were much faster to text classification. Theoretically, although deep neural networks suggest higher representational power than shallow models, but it is not clear if simple text classification problems.

Using the proposed model of consecutive deep learning framework, the pre-processing and classification of side effects were used for the three classes of drugs.

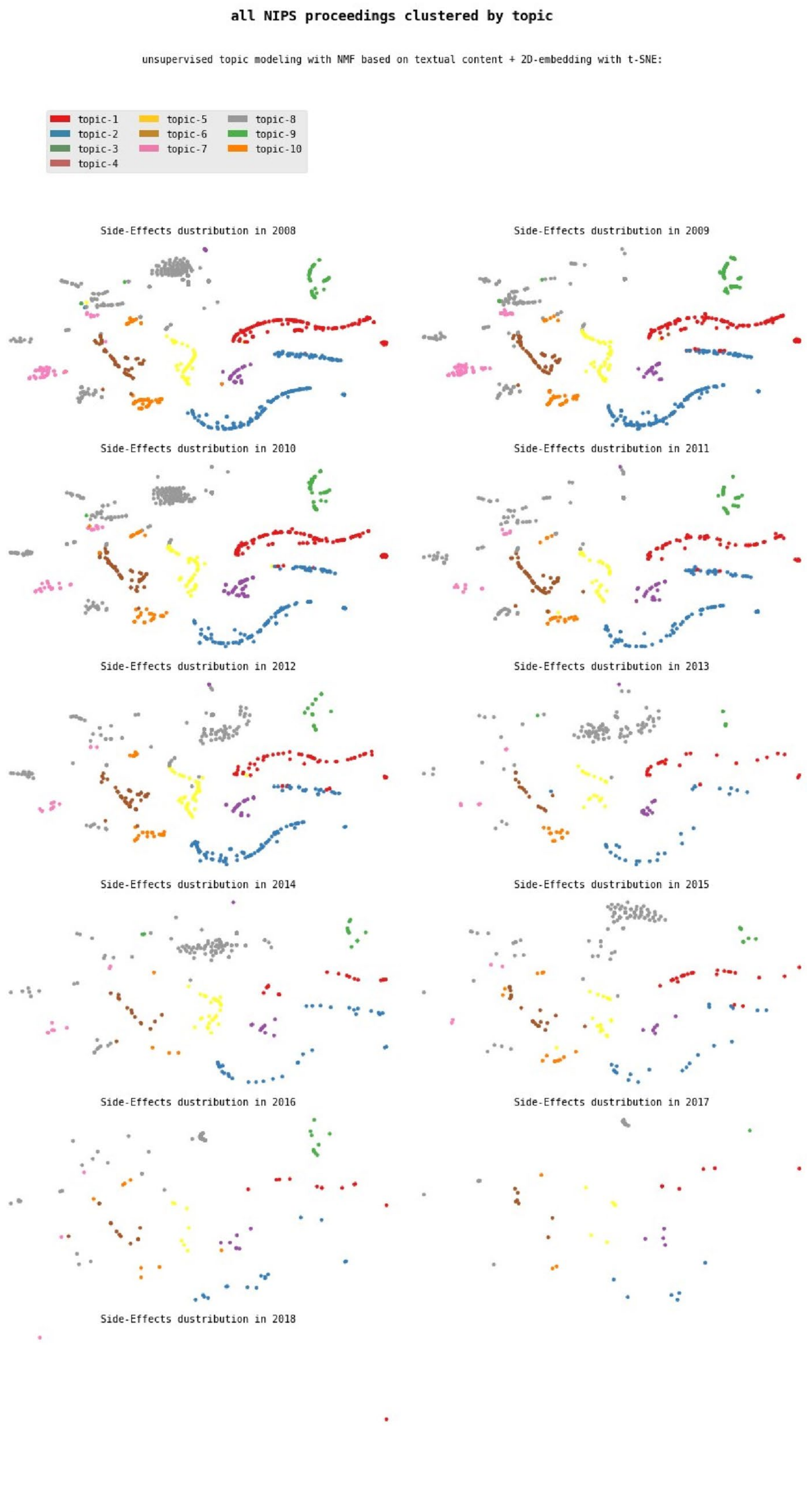


Fig. 5 Scatterplot of anti-pregnancy medicines topics on the website of *Ask a patient* based on year

Fig. 6 Digestion topic modeling visualization (digestion topic modeling (*Ask a patient)* is reported in Table 12 in Appendix)

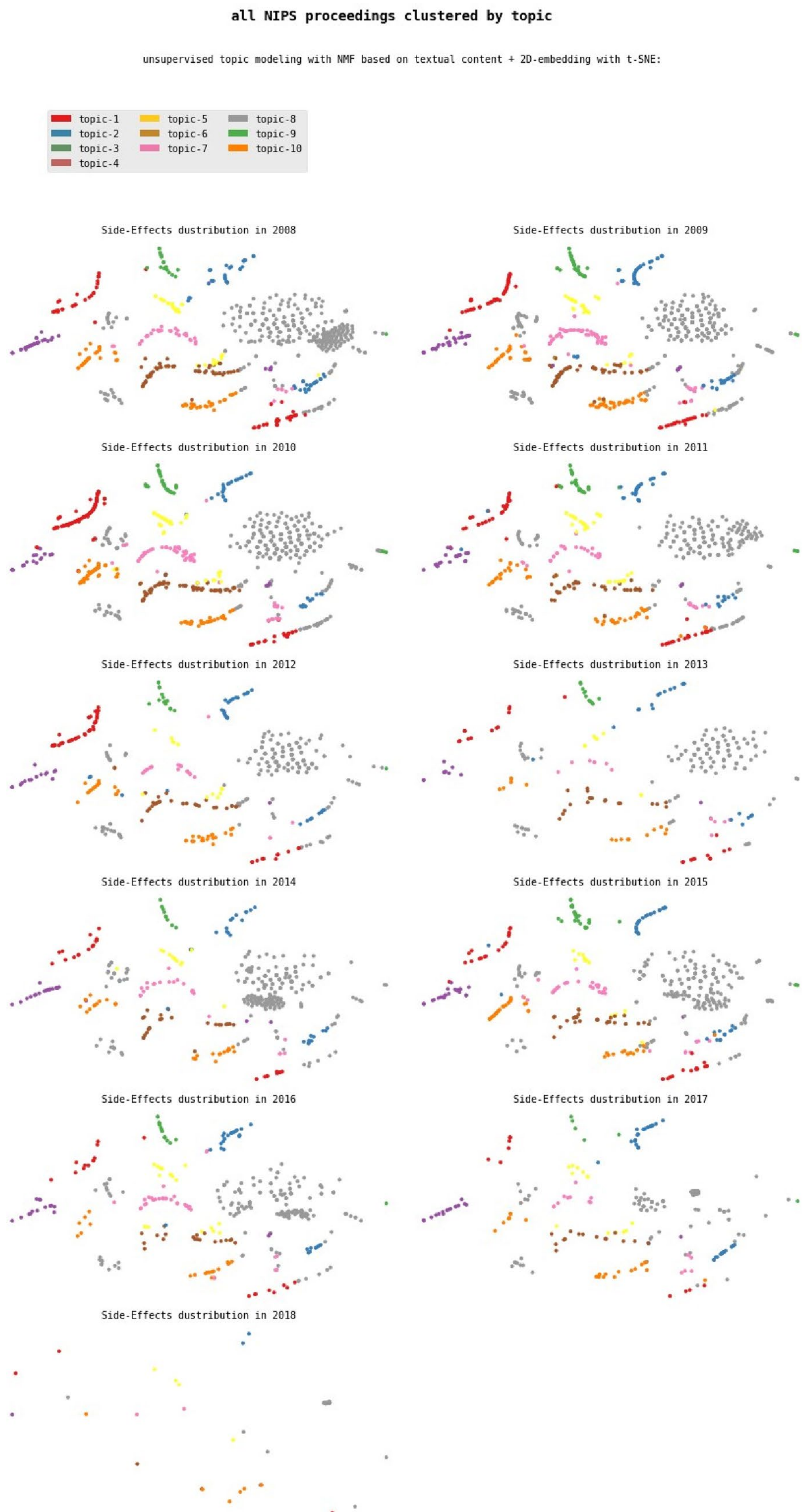
In this work, we used a simple method for text classifications by deep learning models. In contrast to unsupervised trained word vectors derived from word2vec, our word features can be averaged together to generate appropriate sentence representations. In comparison with recent deep learning-based methods, the FastText and HAN were much faster to text classification. Theoretically, although deep neural networks suggest higher representational power than shallow models, but it is not clear if simple text classification problems.

Additionally, in contrast to previous studies, we suggested an end-to-end solution based on deep learning models that do not require any handcrafted features and data pre-processing. Our experimental findings show that each model significantly outperforms baseline methods for different datasets.

8 Conclusions

The users' comments on identifying the side effects of drugs presented in a website, namely *Ask a patient*, were investigated, then a combined classification based on three types of diseases which were mostly commented on were extracted. Through analysis of the data using deep learning method, it was found that users' comments on side effects of drugs were biased, as their comments were not to be evaluated, and it was voluntary. The comments were classified using topic modeling, then some reports similar to the reports issued by Sider and WebMD were issued; however, the reports were different in frequency. As a case in the point, the side effects had been reported with a high frequency in Sider and WebMD, while typical users did not report those side effects very much. On the other hand, some other side effects not reported by Sider and WebMD had attracted typical users' attention. Our findings enable the efficient use of vast batch sizes, significantly reducing the number of parameter updates required to train a model. This has the potential to dramatically reduce model training times. To sum up, using the data from social media in studies on social media opens a wide and novel window in the field of drug studies. The results show that the data from social media may have noise, or may not be reliable. Accordingly, social media can be employed as a secondary source in identifying the side effects of drugs rather than a substitution for traditional and scientific methods of identifying side effects. The option of reporting 'unregistered side effect' shows that a great deal of data, which have not been reported in drug studies, can be extracted from social media. The side effects may have appeared due to the dosage or procedure for using the drug, or it may have been appeared due to interference from other drugs. The model proposed in this study can be used for immediate identification of pharmacological events which most probably leads to immediate reaction and on-time discovery to these events.

Fig. 7 Scatterplot of digestion medicines topics on the website of *Ask a patient* based on year



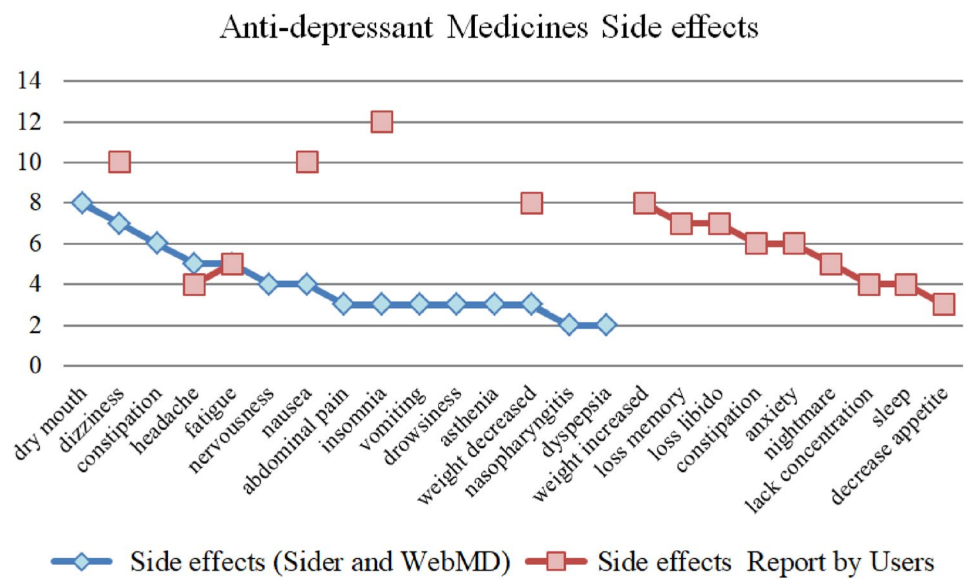


Fig. 8 Comparison of topic modeling of users' comments with the side effects reported on the websites of Sider and WebMD (Neurotic drugs)

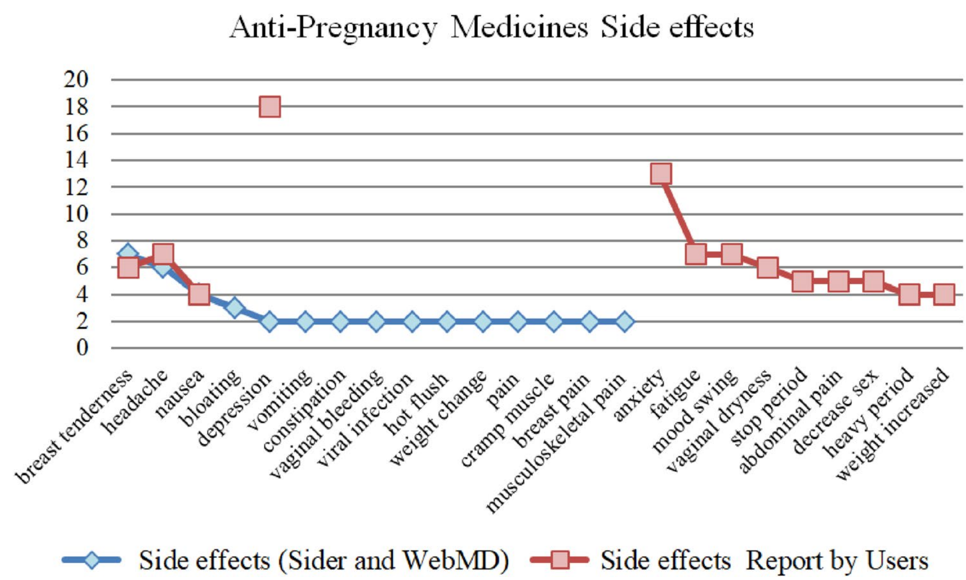


Fig. 9 Comparison of topic modeling of users' comments with the side effects reported on the websites of Sider and WebMD (Anti-pregnancy drugs)

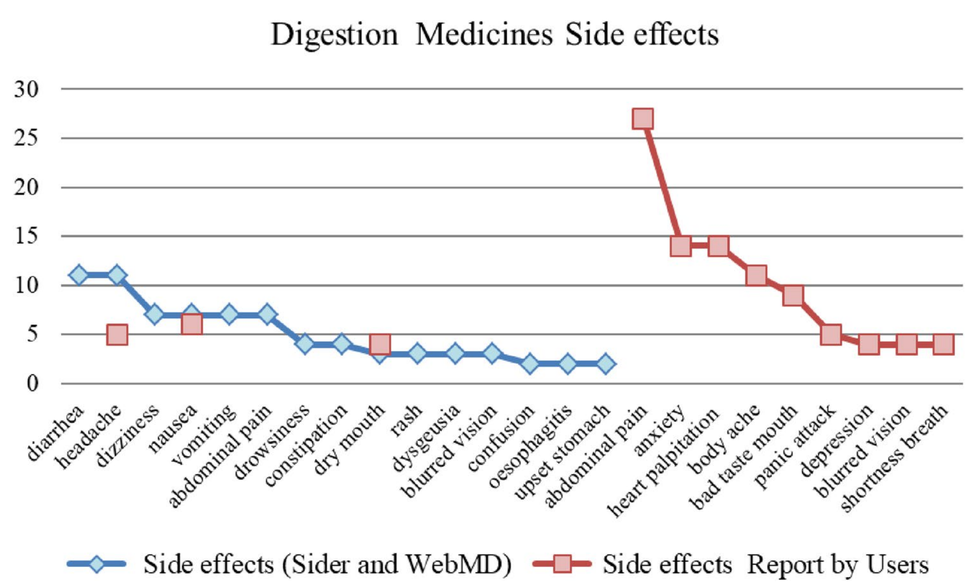


Fig. 10 Comparison of topic modeling of users' comments with the side effects reported in the websites of Sider and WebMD (Digestion drugs)

Appendix

See Figs. 11, 12, 13; Tables 7, 8, 9, 10, 11, and 12.

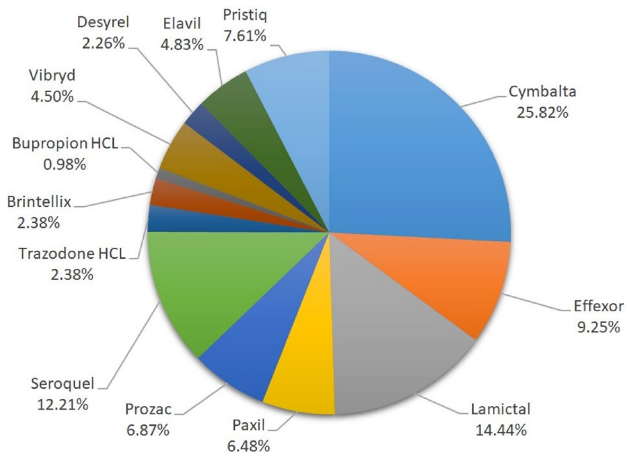


Fig. 11 Anti-depressant medicines side effects (4929 Comments)

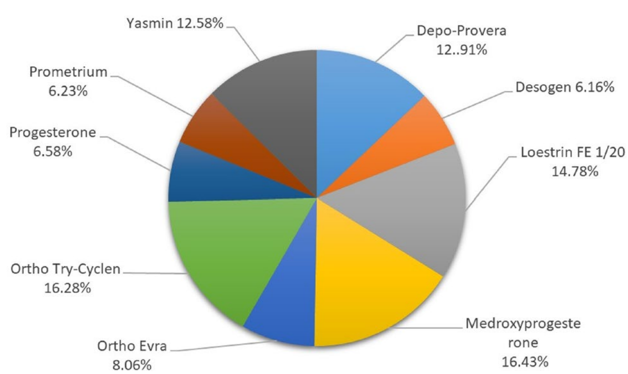


Fig. 12 Anti-pregnancy medicines side effects (4149 comments)

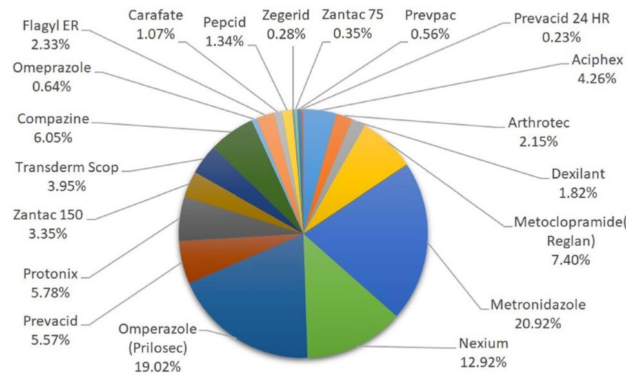


Fig. 13 Digestion medicines side effects (3995 comments)

Table 7 Distribution of data in cross-validation phase

Train	Test	Validation
Neurotic and anti-depression medicines		
4437	492	984
Anti-pregnancy medicines		
3735	414	828
Digestion medicines		
3596	399	798

Table 8 Confusion Matrix – HAN (128-0.001)

Predicted actual	0	1	2	All
0	468	7	17	492
1	19	387	8	414
2	35	13	351	399
All	522	407	376	1305

Table 9 Confusion matrixes—FastText (128-0.001)

Predicted actual	0	1	2	All
0	451	20	21	492
1	15	388	11	414
2	31	20	348	399
All	497	428	380	1305

Table 10 Anti-depressant topic modeling (*Ask a patient*)

Topic #0:	Topic #1:	Topic #2:	Topic #3:	Topic #4:	Topic #5:	Topic #6:	Topic #7:	Topic #8:	Topic #9:
weight gain	dry mouth	memory loss	vivid dream	hair loss	loss appetite	brain zap	weight loss	panic attack	miss dose
extreme weight	mouth constipation	severe memory	night vivid	loss weight	nausea loss	loss libido	appetite weight	suicidal thought	dose hour
increase appetite	mouth weight	loss confusion	dream nightmare	loss memory	loss libido	zap dizziness	decrease appetite	mood swing	dose dizzy
major weight	blur vision	loss trouble	insomnia vivid	blur vision	appetite weight	dizziness brain	slight weight	anxiety panic	withdrawal symptom
massive weight	gain dry	loss weight	dream night	loss hair	mouth loss	horrible brain	loss loss	increase anxiety	dizziness miss
gain–loss	nausea dry	loss loss	decrease libido	joint pain	insomnia loss	depression anxiety	nausea weight	restless leg	hour miss
constipation weight	mouth headache	loss memory	gain vivid	gain hair	headache loss	inability orgasm	week weight	weird dream	nausea dizziness
gain increase	headache dry	blur vision	dream increase	loss insomnia	taste mouth	sleep paralysis	loss weight	depression suicidal	zap miss
gain constipation	mouth sleepiness	long memory	increase dose	memory problem	increase depression	withdrawal symptom	gain weight	extreme fatigue	dose miss
rapid weight	mouth loss	gain memory	dream decrease	muscle ache	dizziness loss	zap dose	loss severe	severe panic	24 h
gain fatigue	constipation dry	confusion memory	acid reflux	itchy scalp	trouble sleep	flu symptom	insomnia weight	lack emotion	dose day
lose weight	extreme dry	dizziness memory	dose vivid	week stop	upset stomach	zap miss	loss increase	anti-depressant	headache nausea
loss libido	mouth week	slight memory	extremely vivid	extreme weight	appetite day	horrible withdrawal	loss decrease	start medication	dose vivid
gain weight	mouth dizziness	loss libido	sleep vivid	memory impairment	loss sex	zap severe	loss month	leg syndrome	dose brain
fatigue weight	mouth insomnia	loss hair	heart palpitation	dry skin	nausea vomit	dose miss	brain fog	night terror	dose night
mouth weight	mouth night	brain fog	lose weight	loss dry	day nausea	zap nausea	headache weight	trouble sleep	gain weight
slight weight	blood pressure	lack concentration	day night	make sense	fatigue loss	gain brain	hour sleep	anxiety depression	depression anxiety
gain month	appetite dry	night sleep	sleep day	vivid nightmare	increase anxiety	nausea brain	loss nausea	increase suicidal	pin needle
gain dry	sleep dry	slur speech	night loss	constipation fatigue	appetite loss	nausea constipation	loss sleep	heart race	severe withdrawal
loss sex	ring ear	mood swing	day sleep	nausea dizziness	stomach pain	extreme dizziness	delay ejaculation	anxiety increase	electric shock

Table 11 Anti-pregnancy topic modeling (*Ask a patient*)

Topic #0:	Topic #1:	Topic #2:	Topic #3:	Topic #4:	Topic #5:	Topic #6:	Topic #7:	Topic #8:	Topic #9:
weight gain	mood swing	breast tenderness	hot flash	hair loss	birth control	panic attack	loss sex	sore breast	weight loss
gain depression	swing depression	nausea breast	flash night	loss weight	gain weight	depression anxiety	gain–loss	abdominal pain	clear skin
slight weight	severe mood	extreme breast	day hot	loss appetite	blood clot	anxiety panic	swing loss	gain sore	light period
gain mood	extreme mood	slight breast	night hot	gain hair	tri cyclen	severe anxiety	fatigue loss	breast acne	period weight
swing weight	depression mood	swing breast	swing hot	anxiety depression	lose weight	severe depression	vaginal dryness	breast nausea	loss period
depression weight	swing weight	tenderness weight	low pain	depression hair	ortho tri	attack depression	sex depression	lose weight	loss loss
bloat weight	gain mood	tenderness mood	vivid dream	dry eye	control pill	attack anxiety	anxiety loss	zero sex	loss acne
gain acne	bad mood	increase appetite	light head	extreme hair	recommend birth	heart palpitation	total loss	cramp mood	increase sex
yeast infection	headache mood	severe breast	depo shot	joint pain	ortho evra	depression panic	moodiness loss	extreme fatigue	loss fatigue
extreme weight	horrible mood	tenderness nausea	depo provera	vaginal dryness	period month	suicidal thought	sex fatigue	vaginal dryness	yeast infection
gain anxiety	vaginal dryness	tenderness headache	long period	swing hair	month stop	anxiety depression	depression loss	chest pain	regular period
gain sex	swing anxiety	tenderness depression	pill day	heart palpitation	stop period	extreme anxiety	sex weight	month period	lot weight
decrease sex	major mood	tenderness increase	severe cramp	loss loss	month period	severe panic	sex mood	breast nipple	loss appetite
headache weight	swing headache	headache breast	headache nausea	heavy period	sick stomach	chest pain	sex vaginal	dry mouth	skin weight
increase appetite	anxiety mood	cramp breast	race heart	sex hair	start pill	swing depression	dryness loss	start period	period cramp
gain increase	increase appetite	tenderness loss	trouble sleep	loss acne	period heavy	extreme depression	extreme fatigue	vivid dream	decrease appetite
gain weight	swing irritability	light period	heart attack	loss extreme	make gain	anxiety weight	headache loss	fluid retention	severe depression
low sex	fatigue mood	tenderness swell	gain bloat	loss depression	heavy period	brain fog	loss loss	breast cramp	vaginal dryness
gain–loss	nausea mood	miss period	fatigue mood	severe depression	blood thinner	swing anxiety	race heart	breast mood	appetite weight
gain moodiness	swing sex	gain breast	anxiety insomnia	painful intercourse	body use	headache anxiety	painful intercourse	day provera	loss libido

Table 12 Digestion topic modeling (*Ask a patient*)

Topic #0:	Topic #1:	Topic #2:	Topic #3:	Topic #4:	Topic #5:	Topic #6:	Topic #7:	Topic #8:	Topic #9:
taste mouth	panic attack	dry mouth	stomach pain	joint pain	heart palpitation	chest pain	anxiety depression	blur vision	stomach cramp
metallic taste	anxiety panic	extreme dry	severe stomach	muscle pain	anxiety heart	blood pressure	severe anxiety	dizziness blu	severe stomach
dark urine	extreme anxiety	mouth headache	pain nausea	pain muscle	shortness breath	shortness breath	loss appetite	mouth blur	cramp pain
bad taste	depression anxiety	extremely dry	pain stomach	weight gain	blood pressure	pain chest	shortness breath	pain blur	cramp nausea
loss appetite	severe panic	mouth bad	bad stomach	pain joint	palpitation anxiety	pain anxiety	mood swing	vision blur	cramp diarrhea
metal taste	race heart	severe dry	pain cramp	muscle joint	hair loss	anxiety chest	depression fatigue	weight gain	nausea stomach
horrible taste	crawl skin	blurry vision	pain constipation	severe joint	brain fog	pain heart	extreme anxiety	fatigue blur	nausea vomit
nasty taste	suicidal thought	headache dry	pain bad	muscle weakness	palpitation dizziness	high blood	weight loss	sensitivity light	headache stomach
mood swing	attack anxiety	patch day	bloat stomach	brain fog	high blood	heart attack	nausea loss	extremely dry	diarrhea stomach
loose stool	think die	bad taste	headache stomach	pain pain	tightness chest	hand foot	depression loss	poor concentration	loose stool
flu symptom	severe anxiety	mouth dry	sore throat	pain severe	muscle twitch	weight gain	muscle spasm	sore throat	brain fog
horrible metallic	brain fog	dizziness dry	mouth stomach	severe headache	dizziness heart	palpitation chest	brain fog	vision anxiety	muscle cramp
light head	attack depression	mouth blur	pain bloat	body ache	headache heart	muscle pain	suicidal thought	fog blur	cramp bloat
bitter taste	heart race	mouth loss	pain anxiety	ring ear	lump throat	pain tightness	depression panic	remove patch	dark urine
upset stomach	heart rate	brain fog	pain headache	pain shoulder	anxiety attack	hair loss	sore throat	mood swing	bad stomach
mouth dark	shortness breath	light head	pain severe	leg cramp	light headedness	tightness chest	extreme fatigue	headache dizziness	cramp stomach
day day	hand foot	wear patch	diarrhea stomach	pain fatigue	trouble sleep	pain shortness	trouble sleep	extreme dry	sick stomach
extreme nausea	horrible anxiety	abdominal cramp	terrible stomach	pain swell	pain heart	race heart	ring ear	mental fog	diarrhea nausea
extreme fatigue	lose mind	muscle cramp	pain day	pain leg	light head	heart rate	major anxiety	48 h	dizziness stomach
metalic taste	horrible panic	mouth throat	body ache	blurry vision	race heart	rapid heartbeat	race heart	weight loss	cramp severe

References

Akhtyamova L, Alexandrov M, Cardiff J (2017a) Adverse drug extraction in twitter data using convolutional neural network. In: 28th international workshop on database and expert systems applications (DEXA). Lyon, France. https://doi.org/10.1109/DEXA.2017.34

Akhtyamova L, Ignatov A, Cardiff J (2017b) A large-scale CNN ensemble for medication safety analysis. arXiv: 1706.0554

Ask a Patient (2001) A website for research drugs and health care topics. https://www.AskaPatient.com/

Bordet R, Gautier S, Louet HL, Dupuis B, Caron J (2001) Analysis of the direct cost of adverse drug reactions in hospitalised patients. Eur J Clin Pharmacol 56:935–941

Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP (1997) Adverse drug events in hospitalized patients. excess length of stay, extra costs, and attributable mortality. JAMA 277:301–306. https://doi.org/10.1001/jama.1997.03540280039031

Ginn R, Pimpalkhute P, Nikfarjam A, Patki A (2014) Mining twitter for adverse drug reaction mentions: a corpus and classification benchmark. Biotxtm

Joulin A, Grave E, Bojanowski P, Mikolov T (2017) Bag of tricks for efficient text classification. In: 15th Conference of the European chapter of the association for computational Linguisti, vol 2, pp 427–431

Kongkaew C, Noyce PR, Ashcroft DM (2008) Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. Ann Pharmacother 42:7–8. https://doi.org/10.1345/aph.1L037

Lee K, Qadir A, Hasan SA, Prakash A, Liu J, Farri O (2017) Adverse drug event detection in tweets with semi-supervised convolutional neural networks. Geneva, Switzerland, pp 705–714. http://dx.doi.org/10.1145/3038912.3052671

Sarker A, Gonzalez G (2015) Portable automatic text classification for adverse drug reaction detection via multi-corpus training. J Biomed Inform 53:196–207. https://doi.org/10.1016/j.jbi.2014.11.002

Sultana J, Cutroneo P, Trifirò G (2013) Clinical and economic burden of adverse drug reactions. J Pharmacol Pharmacother. https://doi.org/10.4103/0976-500X.120957

Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E (2016) Hierarchical attention networks for document classification, pp 1480–1489