Applied Computer Science, vol. 16, no. 1, pp. 41–59
doi:10.23743/acs-2020-04

Submitted: 2020-02-01
Revised: 2020-02-19
Accepted: 2020-03-13

Deep Learning, topic modeling, Text Mining, ADR, NMF

Behnaz ESLAMI[*], *Mehdi HABIBZADEH MOTLAGH*[**],
Zahra REZAEI[***], *Mohammad ESLAMI*[****],
Mohammad AMIN AMINI[*****]

UNSUPERVISED DYNAMIC TOPIC MODEL FOR EXTRACTING ADVERSE DRUG REACTION FROM HEALTH FORUMS

Abstract

The relationship between drug and its side effects has been outlined in two websites: Sider and WebMD. The aim of this study was to find the association between drug and its side effects. We compared the reports of typical users of a web site called: "Ask a patient" website with reported drug side effects in reference sites such as Sider and WebMD. In addition, the typical users' comments on highly-commented drugs (Neurotic drugs, Anti-Pregnancy drugs and Gastrointestinal drugs) were analyzed, using deep learning method. To this end, typical users' comments on drugs' side effects, during last decades, were collected from the website "Ask a patient". Then, the data on drugs were classified based on deep learning model (HAN) and the drugs' side effect. And the main topics of side effects for each group of drugs were identified and reported, through Sider and WebMD websites. Our model demonstrates its ability to accurately describe and label side effects in a temporal text corpus by a deep learning classifier which is shown to be an effective method to precisely discover the association between drugs and their side effects. Moreover, this model has the capability to immediately locate information in reference sites to recognize the side effect of new drugs, applicable for drug companies. This study suggests that the sensitivity of internet users and the diverse scientific findings are for the benefit of distinct detection of adverse effects of drugs, and deep learning would facilitate it.

[*] Islamic Azad University, Science and Research Branch, Department of Computer Engineering, Islamic Azad University, Tehran, Iran, behnazeslami30@gmail.com
[**] P/S/L Group, 1801 McGill College Ave, Montreal, Quebec H3A 2N4, Montreal, Canada
[***] University of Kashan, Department of Computer and Electrical Engineering, Isfahan Province, Qotb-e Ravandi Blvd, Kashan, Iran
[****] Islamic Azad University of Qazvin,Department of Computer Engineering, Qazvin, Iran
[*****] Islamic Azad University of Jasb, Department of Computer Engineering, Markazi, Iran

1. INTRODUCTION

The Adverse Drug Reaction (ADR) is defined as “an undesirable effect”. The ‘side effect’ does not have the exact terminology for inadvertent and secondary effect, observed during therapy. In fact, the interpretation of term “side effect” may vary between two different individuals. However, adverse drug reactions could be considered as the result of toxicity from all kinds of drugs. Apparently, 3 to 7% of all hospitalizations have been due to adverse drug reactions (Kongkaew, Noyce & Ashcroft, 2008). And ADRs noticeably increase patient’s hospitality costs (Sultana, Cutroneo & Trifirò, 2013; Miranda, 2018). According to the annual report of the Agency for Healthcare Research and Quality, over 770,000 patients were injured and/or died in hospitals due to adverse drug reactions in each year (Rison, 2013).

Based on similar singling pathways and cellular structures, involved in normal or abnormal conditions, the same expectation on side effect and actual treatment effect would probably make the uniform pattern for medication. The goal of any drug administration needs to focus on differentiation between negative and positive effect of targeted drug as much as possible, which is required to be tested case by case. The focus of our study is to investigate into appropriate dosage of drugs, since the biological response of each individual to different medication may be various, i.e. one specific drug probably has unexpected destructive effect on one individual, while it is safe for others, thus the interaction between drug and cells need to be adjusted, whose index is normalization of drug dosages per case. Fortunately, there have been available reports for drug interaction in social media which help public have good understanding of side effect. For instance, it has been reported that aspirin and warfarin interfere with clot formation in blood vessels and the subsequently bleeding time would take longer. Another example is the feedback of food or herbs to drugs which modifies their effects, i.e. it has been reported that the level of cholesterol in the circulatory system is reduced by statins however, high fat diets have an opposite effect on blood cholesterol level. Also, St. John’s Wort could make bipolar individual hyperactive in spite of consumption of the antidepressant drug (Bordet, Gautier, Louet, Dupuis & Caron, 2001).

It takes a well-trained reader a lot of time to screen ADRs by looking through relevant literatures without using a machine reader. Therefore, it is crucially valuable for experts to benefit from automated system in order to find ADRs in publications as fast and efficiently as possible (Classen, Pestotnik, Evans, Lloyd & Burke, 1997). The detection of ADRs have not been initially well-structured and just obtained through communication between health professionals and patients or published case reports, available in MEDLINE, PubMed or other publicly available datasets (Rison, 2013; Vallano et al., 2005). Hence, society needs an alternative approach to detect side effects of the clinical medications. The social media is capable of producing novel and reliable data sources for the side effects of drugs.

In fact, through the social media, special events in the field of health could be identified and managed. "Ask a patient" is the web page that allows patients to share and compare medication experiences, and was granted Webby Award for the best website in the Pharmaceutical Category in 2012. The "Ask a patient" database contains more than 4,000 chemically prepared and prescribed drugs, approved by FDA's Center for Drug Evaluation and Research.

Comments over prescription or the counter drugs, found in this web page, would be based on fine-tuned search criteria (age, gender, symptom, etc.). However, the difference between written and oral language in social media creates some noises. Also, lack of a suitable structure and imbalance data in drug groups are considered as important challenges in classification of data, retrieved from social media. Accordingly, in spite of richness of health-related data in social media, it seems not to be practical to use this type of data for the purpose of ADR detection.

In this study, we identify drug side effect based on three main criteria:

1. An automated deep learning was applied to extract features from social media. The comments of "Ask a patient" website's users, were processed to describe side effects and thus reduce the difference between written and oral language and dampen down the noise effect.
2. The efficacy of deep learning method in classification of data from "Ask a patient" was approved by the quality of the outcome. The results showed that deep learning performance benefits from high accuracy and speed, simultaneously.
3. Advantage and disadvantage of each comment were compared with those of already reported ones in Sider and WebMD web pages. In order to achieve that, deep learning method HAN (Yang et al., 2016) was employed to classify users' comments. Then, the non-monitoring method (NMF) of topic modeling was administered to determine specific topics in each group of drugs.

2. RELATED WORKS

Some studies have hitherto investigated into the side effect of drugs using social media as tool. For example, Sarker and Gonzalez highlighted the importance of combined usage of advanced NLP-based information generation and traditional text classification (Support Vector Machine, Naïve Bayes and Maximum Entropy) to accurately detect and classify sentences concerning ADR (Sarker & Gonzalez, 2015). Aligned with that, Ho et al. suggested the automated detection of data related to ADR by searching relevant database; they prepared a systematic review and concise information about suitable approach to envisage ADEs, pointed out in social media (Ho, Le, Thai & Taewijit, 2016).

Also, Ginn and coworkers applied two supervised machine learning approaches (NB and SVM) on a wide range of annotated medications in association with ADR tweets (Ginn et al., 2014). Although, the classifier showed moderate performance, it was considered as the base for future development in advanced techniques. Aligned with this approach, they used Convolutional Neural Networks (CNN) model, which applied word2vec embedding for classification of Twitter comments. In contrast to other models, their proposed model not only used a small fraction of features for data collection, but also showed high performance in text classification procedures (Akhtyamova, Alexandrov & Cardiff, 2017a). Recent attempts have been made to benefit from specific type of deep learning to enhance quality of ADR discovering through extraction of sentences and entities, available in social media. Gupta et al. suggested a two-step method to extract pointed out adverse event, i.e. it initially predicts drug with regard to input contexts, unsupervisedly, and then it repeats same direction in a supervised way (Gupta, Pawar, Ramrakhiyani, Palshikar & Varma, 2018). In parallel, Tan et al. offered the summary of data base and automated systems to support ADRs detection (Tan et al., 2016). Also, Harpaz et al. presented the synopsis on using text mining for the purpose of Adverse Drug Events (ADEs) detection, in publicly available literature or web pages (Harpaz et al., 2014).

In addition, Lee and colleagues put forward a semi-supervised CNN-based framework to classify the adverse drug event (ADE) in Twitter. A Twitter dataset was used in PSB 2016 Social Media Shared Task, leading to high performance classification of ADE with 9.9% F1-Score (Lee et al., 2017). It is good to be pointed out that ADE detection surveillance systems require small number of labeled instances. Also, Akhtyamova et al, presented a CNN-based architecture, composed of numerous parameters to predict adverse drug reaction based on the quantity of votes (Akhtyamova, Alexandrov & Cardiff, 2017b). They utilized a large scale of medical dataset, derived from medical websites, in order to evaluate the mode of performance. In contrast to previously reported networks, the proposed end-to-end model does not require handcrafted features and data pre-processing, and it resulted in an enormous improvement in standard CNN based methods.

Finally Rezaei et al, suggested three methods for preprocessing of data analyses and used numerous deep learning methods for text classification. Compared to current deep learning-based networks, their results showed that the FastText, CNN, and HAN were much faster and more accurate. According to deep learning models, they suggested the approach of end-to-end, in which artificial attribute and preprocessed information are not necessary. The obtained results demonstrated that the proposed models would significantly improve the performance of baseline methods for different datasets. They noticed that increasing batch size during training steps considerably reduced the learning rate in the network. Conversely, they tested various optimizers including SGD, RMS, and Adam in their custom datasets, and found that Adam shows better results compared to RMS and SGD (Rezaei, Ebrahimpour-Komleh, Eslami, Chavoshinejad & Totonchi, 2020).

This study aims to investigate the written topic modeling of typical users and identify the changes in comments, which have been reported from 10 years ago. We designed a model that provides researchers with immediate capability of analyzing comments through combined deep learning methods.

3. METHOD

This paper is organized into two sections; classification and extraction of topics (Fig. 1).

3.1. Classification

3.1.1. Data Sources

Prior to data collection, we selected a set of interesting drugs, which were likely to have a large number of associated comments in "Ask a patient" database. We chose drugs that were prescribed for chronic diseases and syndromes, i.e. the medication with high prevalent prescription and referred comments. The names of the medications were reported in separate classes (Anti-depressant drugs, Anti-Pregnancy drugs and Gastrointestinal drugs) in figures 2 to 4.

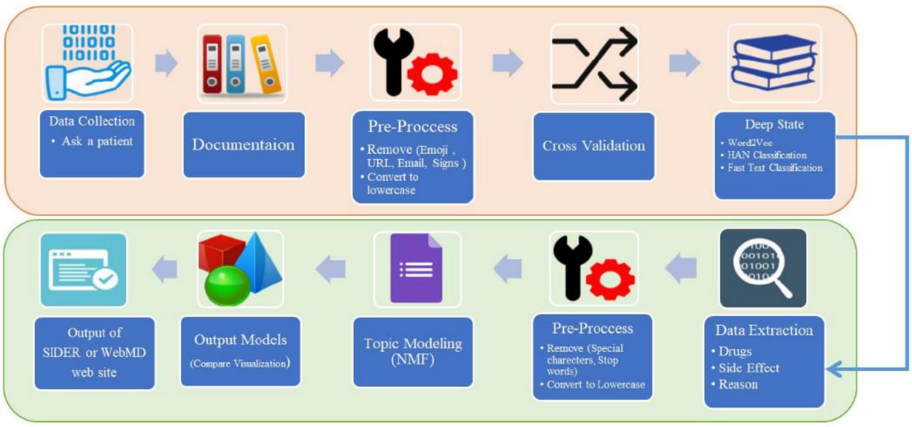

Fig. 1. The workflow of the proposed deep learning based strategy is illustrated

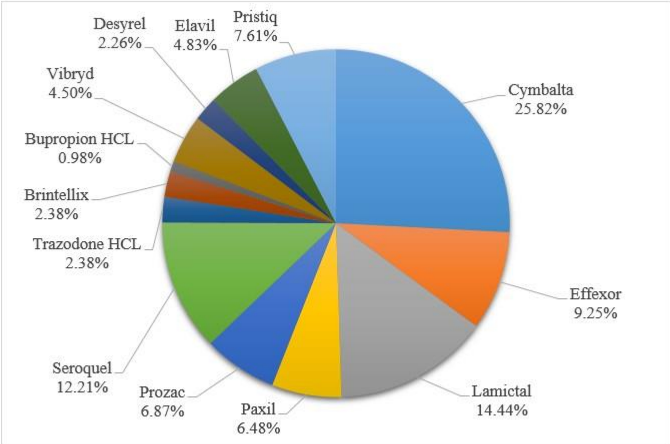

Fig. 2. Anti-Depressant Medicines Side effects (4929 Comments)

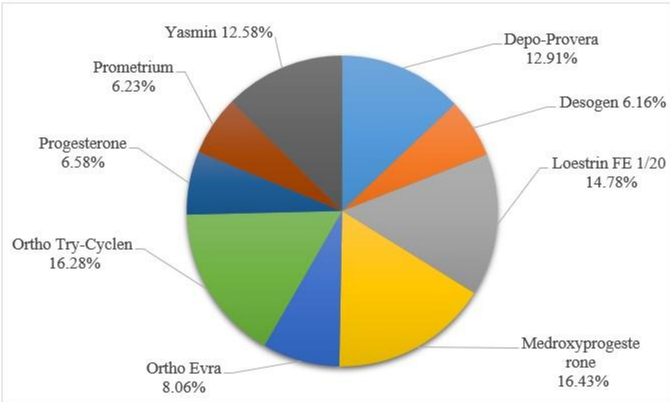

Fig. 3. Anti-Pregnancy Medicines Side effects (4149 Comments)

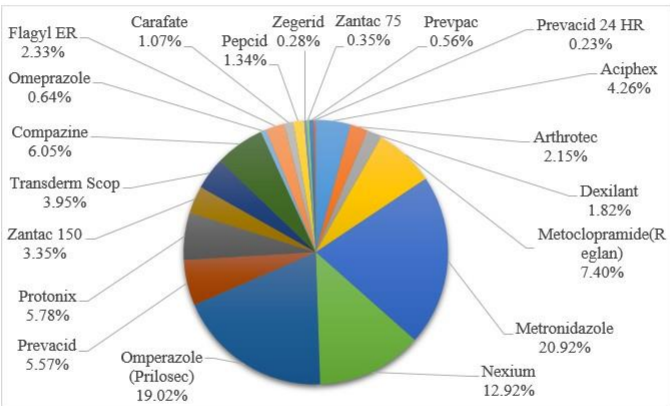

Fig. 4. Digestion Medicines Side effects (3995 Comments)

3.1.2. Preprocessing

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

- Data shuffling,
- Converting all uppercase words to lowercase ones,
- Elimination of special characters like: @, !, /, *, $ and etc.,
- Removal of stop word: at, of, the, … ,
- Correction of words with repeated characters like: pleaseeeeeeeee and/or yessss,
- Conversion of contractions to base format like: I'm → I **am**,
- Lemmatization: I started taking almost two months ago. → I **start take** almost two months ago.

3.1.3 Cross Validation

In order to achieve the best performance with regard to new data, we wished to find the appropriate values of the complexity parameters, leading to optimal model. If the amount of data was high, the procedure would have been divided into three subsets; the training, the validation and the test sets. Among the diverse complex models that have been trained, we selected the one that had the best predictive and effective performance, and was confirmed by the data in the validation set. However, the data supply was limited for training and test set, which led to the increase of the generalized error. Thus, cross validation was applied to reduce these types of error and prevent over-fitting. The data distribution for each group is shown in Table 1.

Tab. 1. Distribution of data in Cross-Validation phase

Medicines Category	Train Phase Docs	Test Phase Docs	Validation Phase Docs
Neurotic and Anti Depression Medicines	4437	492	982
Anti-pregnancy Medicines	3735	414	828
Digestion Medicines	3596	399	798

3.1.4. Deep Classification

The applied methods for data classification are HNN (Yang et al., 2016) and FastText (Joulin, Grave, Bojanowski & Mikolov, 2016) with similar word2vec section. Once word2vec generated, this file would be used for further investigations.

3.1.4.1. HAN Method

Hierarchical Attention Network (HAN) has two distinctive characteristics: (I) a hierarchical structure and documents, (II) two-phase mechanism of attention, which enables HAN to differentially put words or sentences next to each other within the structure of the document. In addition to these two characteristics, HAN network is composed of quite a few parts including, i.e. a word sequence encoder, a word-level attention layer, a sentence encoder and a sentence-level attention layer. HAN works based on a positive role of sentences and document structure in modeling.

3.1.4.2. FastText Method

This method demonstrates a simple and efficient approach for classification of the texts and its expressions. Large numbers of studies show that the classification of texts with this method is faster in comparison with deep learning methods, with regard to accuracy and applied commands for training and evaluation.

Tab. 2. (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, SGD** and **RMS prob**"}} **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 Data Set** {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 Create a dataset provider that loads data from the dataset // Return [**Content, Label**] Create Training Operations **Run the Training**

In terms of structure, there are two major and influential differences, as follow:

- Softmax: It is a hierarchy, based on the Huffman encoded tree structure that reduces Time Complexity O(Kd) to O(d log k), where K is number of targets and D is dimension of the hidden layer.
- N-gram features: While Bag of words is invariant to word order; it is very expensive to take simplicity into consideration. Instead, we used bag of n-gram as an additional feature to capture some partial information about local word order, which seems to be more efficient in practice (Table 2).

3.1.4.3. Evaluation Metrics

- Precision (positive predictive value) and recall (sensitivity): These metrics are appropriate fraction of retrieved samples from all and relevant instances. Application of these metrics depends on understanding and measuring of relevance.
- Accuracy: This 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 reliable is the classification output is reliable.
- F-measure: This criterion is a combination of call metrics and accuracy and it is used to find if 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 coincident.

Tab. 3. Evaluation metrics formula

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

3.2. Extracted Topics

3.2.1. Data Sources

Three classes of drugs have been consumed between 2008 and 2018 in figures 2 to 4.

3.2.2. Topic Modeling

As a linear algebraic model, Non-negative Matrix Factorization (NMF) includes high-dimensional vectors and low-dimensional image. Vectors are non-negative in NMF like Principal Component Analysis (PCA). Skewing the vectors towards lower-dimensional form in NMF makes the coefficients non-negative.

The two matrices of **W** and **H**, would be obtained through original matrix A, in which **A = WH**. Also, NMF has an inborn clustering property. **A, 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) are elicited from the documents.
- **H (Coefficient Matrix):** the membership weights for the topics in each document.
- **W** and **H** are calculated by optimization of an objective function (like the *EM algorithm*), and updating both **W** and **H**, iteratively, until they are converged (Table 4).

Tab. 4. NMF topic modeling configuration

NMF Topic Modeling
Initializations: **Number of Topics: {10}** **Number of Top Words: {20}** **Configuration of feature extraction by using TfidfVectorizer: {** **Initials: {** **ngram_range:** {(2, 2)}, **Minimum Document Frequency (min_df):** {2}, **Configuration of NMF Topic Modeling Parameters and fit by TfidfVectorizer: {** **components:** {Number of Topics}, **init: {'**Scikit-Learn implementation of NMF (**including NNDSVD initialization)'},** // better for sparseness **}}}** **Run to extracting Topics**

4. RESULT

4.1. Usage Model

In this study, we benefited from user's comments in "Ask a patient" to extract side effects of drugs. In general, the scale of curser that moves over texts in both FastText and HAN methods is called *Pad_Seq_Len* and we considered quantity equal to 150 for that; because, the maximum size of comments is 150 to pay more attention to the length of sentences and semantic conjugation. Moreover, the value of Embedding dim was 100. We evaluated several optimizations such as *Stochastic Gradient Descent, RMS probe* and *Adam*. That *Adam* shows better results (Table 5).

The value of ngram_range was chosen based on the side effects, extracted from Sider or WebMD websites. Other values such as (1, 2), (2, 3) and (3, 3) were determined but (2, 2) was the best choice (Table 6).

Tab. 5. HAN hyper parameters

Pad_Seq_Len	150
Embedding_Dim	100
Drop_Out_Prob	0.5
Loss	Sigmoid
Optimization	Adam

Tab. 6. Evaluation metrics formula

ngram_range	min_df
(2, 2)	2

4.1. Implementation Model in 3.1

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 tables (Table 7 and 8). As shown in these tables, the best result in each method, the learning rate as well as batch size was evaluated. Also, different criteria have been tested for each type of model according to the type of data, which have been obtained in various values. For example, applying HAN method including Batch size of 128 and learning rate of 0.001 on "Ask a patient" dataset and resulting in highest accuracy (0.924) which is highlighted in Table7.

Tab. 7. Output of deep learning classification (HAN Method) on dataset

Dataset	Method	Batch Size	Learning Rate	Accuracy	Kappa	Recall	Precision	F1 Score
Ask a Patient	HAN	32	0.1	0.881	0.821	0.878	0.887	0.881
				0.883	0.842	0.881	0.885	0.882
				0.908	0.862	0.906	0.911	0.907
		64	0.01	0.889	0.833	0.887	0.891	0.888
				0.873	0.808	0.870	0.876	0.872
				0.921	0.881	0.919	0.924	0.921
		128	**0.001**	0.888	0.831	0.885	0.891	0.887
				0.879	0.818	0.879	0.878	0.879
				0.924	**0.885**	**0.921**	**0.926**	**0.923**

Tab. 8. Output of deep learning classification (FastText Method) on dataset

Dataset	Method	Batch Size	Learning Rate	Accuracy	Kappa	Recall	Precision	F1 Score
Ask a Patient	FastText	32	0.1	0.892	0.837	0.888	0.897	0.892
				0.872	0.806	0.866	0.887	0.870
				0.891	0.836	0.888	0.895	0.891
		64	0.01	0.896	0.843	0.894	0.897	0.895
				0.885	0.827	0.884	0.886	0.885
				0.899	0.848	0.898	0.899	0.898
		128	**0.001**	0.876	0.814	0.876	0.876	0.875
				0.895	0.841	0.892	0.896	0.894
				0.909	**0.863**	**0.908**	**0.909**	**0.909**

4.2. Implementation model in 3.2

Considering the output of the previous phase, the three features i.e. Side effects, reason and drug were used. Accordingly, in each class of drugs (neurotic medicines, anti-pregnancy and gastrointestinal), 10 topics with high priority were selected. As shown in tables 9 to 11, topics of each class are verbally similar.

Tab. 9. Anti-depressant Medicines 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 hour
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

Tab. 10. Anti-depressant Medicines 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

Tab. 11. Anti-depressant Medicines 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 hour	dizziness stomach
metalic taste	horrible panic	mouth throat	body ache	blurry vision	race heart	rapid heartbeat	race heart	weight loss	cramp severe

After extraction of these tables, all are mapped with a similar word, and meaningless topics were deleted. Figures 5, 6 and 7 show the frequency of repetition of topic models.

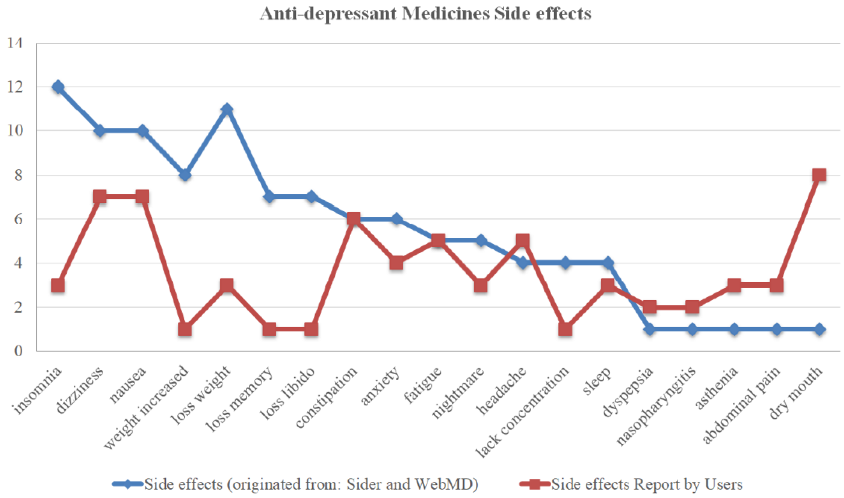

Fig. 5. Comparison of Topic Modeling of users' comments with the side effects reported on the websites of Sider and WebMD (Neurotic drugs)

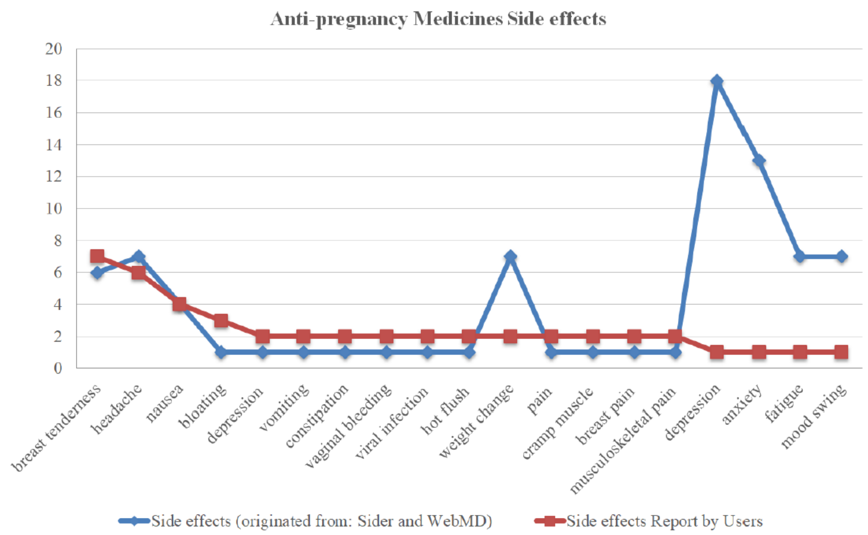

Fig. 6. Comparison of Topic Modelling of users' comments with the side effects reported on the websites of Sider and WebMD (Anti-pregnancy drugs)

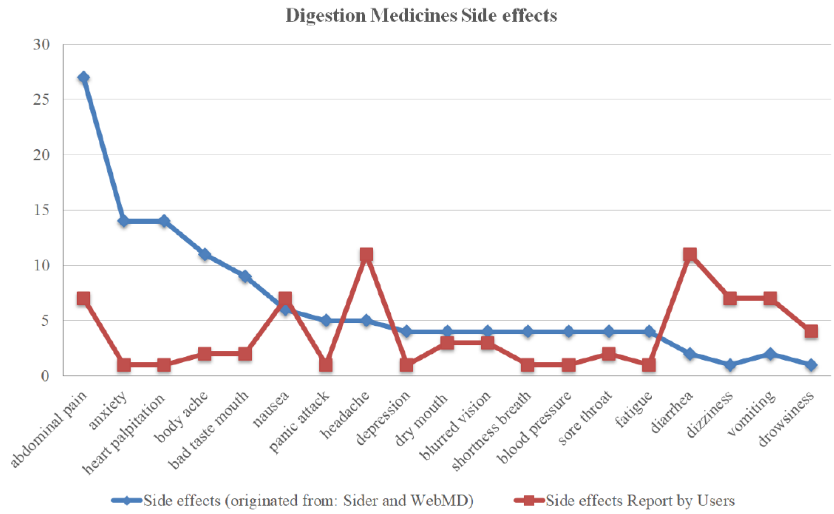

Fig. 7. Comparison of Topic Modeling of users' comments with the side effects reported in the websites of Sider and WebMD (Gastrointestinal drugs)

5. DISCUSSION

In this study, the deep learning methods of HAN and FastText were employed to classify the side effects of three classes of drugs, namely, neurotic, anti-pregnancy and gastrointestinal drugs. The reason for this investigation was high frequency of this drug consumption. Initially, the extracted data from the website "Ask a patient" were introduced to the model. And, in the pre-processing step, special characters, signs and stop words were removed, and other characters were converted into small-case letters in order to improve the text. In next phase, three classes of drugs, the side effect and the association between the former and the latter was investigated. Then, these data were exposed to classification phase (Topic Modelling) to extract 10 topics with high priority from three groups of drugs. The outputs show that the frequency of occurrence of side effects, reported in the comments in "Ask a patient" was different from that in Sider and WebMD.

Finally, the proposed model compared its output on drug's side effects with analyses of report of sites' users. The obtained results of the preliminary analysis of drug classification were presented in confusion matrices and interpreted by taking accuracy rate and false positive ratio into consideration.

In this work, it was found that Fast Text and HAN were much faster for text classification, compared to recent deep learning-based methods. We used a simple method for text classification by deep learning models. In contrast to unsupervised trained word vectors, obtained from word2vec, our word features would approximately generate appropriate sentence representations. Also, in contrast to previous studies, we suggested an end-to-end solution, based on deep learning models which do not need any handcrafted features and data pre-processing.

Our experimental findings show that each model significantly outperforms baseline methods for different datasets. Although deep neural networks, theoretically suggest higher representational power than shallow models, it is still unclear whether simple text classification would create problem or not.

6. CONCLUSION

We investigated the users' comments to identify the side effects of drugs, presented in a website, namely, "Ask a patient", then we extracted combined classification, based on three types of mostly commented diseases. Through analysis of the data with deep learning method, it was found that users' comments on side effects of drugs were biased. On the next step of this study, the comments were classified by Topic Modelling, resulting in some reports, similar to the reports published by Sider and WebMD; however, our reports had different frequency.

Our findings enable us to efficiently and quickly use large size data (batches of sample), and significantly reduce the number of updated parameters that are required for model training.

To sum up, working on publicly available data in social media opens a wide and novel window in the field of drug studies. The results of this study show that the data from social media may have noise, or may not be reliable. Accordingly, social media would be considered as a secondary source to identify side effects of drugs rather than a substitution for traditional and scientific methods of side effect identification. The proposed model in this study is capable of immediate identification of pharmacological events which most likely lead to immediate reaction and on-time discovery of these events.

Acknowledgments

The authors have no proprietary, financial, professional or other personal interest of any nature in any product, service or company. The authors alone are responsible for the content and writing of the paper. There is no conflict of interest in this article.

REFERENCES

Akhtyamova, L., Alexandrov, M., & Cardiff, J. (2017a). Adverse drug extraction in twitter data using convolutional neural network. *In, 2017 28th International Workshop on Database and Expert Systems Applications (DEXA)* (pp. 88–92). Lyon.

Akhtyamova, L., Ignatov, A., & Cardiff, J. (2017b). A Large-scale CNN ensemble for medication safety analysis. In F. Frasincar, A. Ittoo, L. Nguyen & E. Métais (Eds.) *Natural Language Processing and Information Systems. NLDB 2017. Lecture Notes in Computer Science* (vol. 10260, pp. 247–253). Springer, Cham.

Bordet, R., Gautier, S., Louet, H. L., Dupuis, B., & Caron, J. (2001). Analysis of the direct cost of adverse drug reactions in hospitalised patients. *European journal of clinical pharmacology*, *56*(12), 935–941.

Classen, D. C., Pestotnik, S. L., Evans, R. S., Lloyd, J.F., & Burke, J. P. (1997). Adverse drug events in hospitalized patients: excess length of stay, extra costs, and attributable mortality. *Jama*, *277*(4), 301–306.

Ginn, R., Pimpalkhute, P., Nikfarjam, A., Patki, A., O'Connor, K., Sarker, A., Smith, K., & Gonzalez, G. (2014). Mining Twitter for adverse drug reaction mentions, a corpus and classification benchmark. In *Proceedings of the fourth workshop on building and evaluating resources for health and biomedical text processing* (pp. 1–8).

Gupta, S., Pawar, S., Ramrakhiyani, N., Palshikar, G. K., & Varma, V. (2018). Semi-supervised recurrent neural network for adverse drug reaction mention extraction. *BMC bioinformatics*, *19*(8), 212.

Harpaz, R., Callahan, A., Tamang, S., Low, Y., Odgers, D., Finlayson, S., Jung, K., LePendu, P., & Shah, N. H. (2014). Text mining for adverse drug events, the promise, challenges, and state of the art. *Drug safety*, *37*(10), 777–790.

Ho, T. B., Le, L., Thai, D. T., & Taewijit, S. (2016). Data-driven approach to detect and predict adverse drug reactions. *Current pharmaceutical design*, *22*(23), 3498–3526.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers* (pp. 427–431). Association for Computational Linguistics.

Kongkaew, C., Noyce, P. R., & Ashcroft, D.M. (2008). Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. *Annals of Pharmacotherapy*, *42*(7–8), 1017–1025.

Lee, K., Qadir, A., Hasan, S. A., Datla, V., Prakash, A., Liu, J., & Farri, O. (2017). Adverse drug event detection in tweets with semi-supervised convolutional neural networks. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 705–714). Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee. doi:10.1145/3038912.3052671.

Miranda, D. S. (2018). *Automated detection of adverse drug reactions in the biomedical literature using convolutional neural networks and biomedical word embeddings.* SwissText.

Rezaei, Z., Ebrahimpour-Komleh, H., Eslami, B., Chavoshinejad, R., & Totonchi, M. (2020). Adverse Drug Reaction Detection in Social Media by Deepm Learning Methods. *Cell journal*, *22*(3), 319–324.

Rison, R. A. (2013). A guide to writing case reports. *Journal of Medical Case Reports and BioMed Central Research Notes*, *7*, 239. doi:10.1186/1752-1947-7-239

Sarker, A., & Gonzalez, G. (2015). Portable automatic text classification for adverse drug reaction detection via multi-corpus training. *Journal of biomedical informatics*, *53*, 196–207.

Sultana, J., Cutroneo, P., & Trifirò, G. (2013). Clinical and economic burden of adverse drug reactions. *Journal of pharmacology*, *4*(Suppl1), 73.

Tan, Y., Hu, Y., Liu, X., Yin, Z., wen Chen, X., & Liu, M. (2016). Improving drug safety, From adverse drug reaction knowledge discovery to clinical implementation. *Methods*, *110*, 14–25.

Vallano, A., Cereza, G., Pedròs, C., Agustí, A., Danés, I., Aguilera, C., & Arnau, J. M. (2005). Obstacles and solutions for spontaneous reporting of adverse drug reactions in the hospital. *British journal of clinical pharmacology*, *60*(6), 653–658.

Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. (2016). Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics, Human Language Technologies* (pp. 1480–1489). Association for Computational Linguistics.