Computational Linguistics and Intellectual Technologies:
Proceedings of the International Conference "Dialogue 2017"

Moscow, May 31—June 3, 2017

IDENTIFYING DISEASE-RELATED EXPRESSIONS IN REVIEWS USING CONDITIONAL RANDOM FIELDS

Miftahutdinov Z. Sh. (zulfatmi@gmail.com)[1],
Tutubalina E. V. (elvtutubalina@kpfu.ru)[1],
Tropsha A. E. (alex_tropsha@unc.edu)[1,2]

[1]Kazan Federal University, Kazan, Russia
[2]University of North Carolina, Chapel Hill, USA

As the as the volume of user-generated content in social media expands so do the potential benefits of mining social media to learn about patient conditions, drug indications, and beneficial or adverse drug reactions. In this paper, we apply Conditional Random Fields (CRF) model for extracting expressions related to diseases from patient comments. Our method utilizes hand-crafted features including contextual features, dictionaries, cluster-based and distributed word representation generated from unlabeled user posts in social media. We compare our CRF-based approach with deep recurrent neural networks and a dictionary-based approach. We examine different word embeddings generated from unlabeled user posts in social media and scientific literature. We show that CRF outperformed other methods and achieved the F_1-measures of 69.1% and 79.4% on recognition of disease-related expressions in the exact and partial matching exercises, respectively. Qualitative evaluation of disease-related expressions recognized by our feature-rich CRF-based approach demonstrates the variability of reactions from patients with different health conditions.

Key words: disease named entity recognition, opinion expressions, conditional random fields, CRF, information extraction

ВЫЯВЛЕНИЕ СВЯЗАННЫХ С ЗАБОЛЕВАНИЯМИ ВЫРАЖЕНИЙ ИЗ ОТЗЫВОВ ПАЦИЕНТОВ НА ОСНОВЕ МОДЕЛИ УСЛОВНЫХ СЛУЧАЙНЫХ ПОЛЕЙ

Мифтахутдинов З. Ш. (zulfatmi@gmail.com)[1],
Тутубалина Е. В. (elvtutubalina@kpfu.ru)[1],
Тропша А. Э. (alex_tropsha@unc.edu)[1,2]

[1]Казанский федеральный университет, Казань, Россия
[2]Университет Северной Каролины, Чапел-Хилл, США

Ключевые слова: извлечение информации, условные случайные поля, анализ мнений, CRF, извлечение эффектов лекарств, извлечение симптомов заболеваний

1. Introduction

The explosive growth of social media has provided millions of people with the opportunity to share their thoughts or observations related to their health and health care. Repositories of user discussions such as patient portals can be often freely accessed by researchers interesting in social media listening to gather valuable new information about new uses of existing medications, adverse drug reactions, or unknown benefits associated with taking the medications.

A recent trend in text mining research is to move from detecting mentions of genes, gene variants, chemical/drug names, species and other biological concepts towards the broader task of extracting actionable insights from user feedback [1, 14, 17]. Research papers and electronic health records (EHRs) have been the subject of many experimental and clinical studies over the past decade [10]. The task of mining biomedical information from social media instead of articles and EHRs is more challenging due to the informal writing style of a text. Patients who are authors of comments lack formal medical skills to describe observed symptoms and drug reactions as medical concepts. Therefore, there is a growing interest in using machine learning approaches to enhance extraction of medical concepts from social media posts. Applications of these methods include pharmacovigilance and drug repurposing, that focus on extraction of adverse drug reactions (ADRs) and novel drug indications, respectively [7].

Conditional Random Fields (CRF) [16] have been successfully applied to numerous named entity recognition (NER) tasks including recognition of persons and organizations [8, 32], opinion aspects [3, 21], opinion expressions [11], and chemical and medical concepts [12, 18, 20, 23, 34]. In this paper, we apply CRF for the extraction of expressions associated with disease type from social media posts. Disease type consists of (i) entities that specify the reason for taking the drug (e.g., a specific disease name or symptoms of a disease), (ii) outcomes that can be attributed to some action of a drug (e.g., ADRs), and (iii) other findings like patient history. We employ an annotated corpus named CADEC that consists of 1250 medical forum posts taken from AskaPatient.com [14], where each post was manually annotated with mentions of drugs and disease-related entities such as symptoms, ADRs, and clinical findings. We compare CRF with bidirectional Long Short Memory Network (LSTM) and Gated Recurrent Units (GRU) [4, 9] and show that CRF is superior to the alternative approaches. The results of this study suggest that text mining of voluntary patient reports in social media using advanced methods such as CRF could be used as a reliable approach to identifying relationships between diseases (or medical conditions) and drug effects.

2. Related Work

Extr wexplore action of opinion targets (also called aspects) and opinion expressions has been pursued by many researchers using frequency-based methods, unsupervised and supervised methods. Most of the current unsupervised models are based on modifications of Latent Dirichlet Allocation (LDA) [26, 31]. The former are mainly based on Hidden Markov Models [36] and CRF [3, 5, 13]. Recently, bidirectional recurrent neural networks have been shown to outperform CRF on NER tasks [11, 21]. Irsoy and Cardie [11] applied deep Recurrent Neural Networks (RNNs) to extract direct or expressive subjective expressions. Three-layer RNN outperforms CRF, semi-CRF and shallow RNN. Liu et al. [21] applied RNNs for aspect extraction from datasets about laptops and restaurants. RNNs based on pre-trained word embeddings outperformed feature-rich CRF-based models. We mark [29, 30] about active learning and transfer learning as possible directions for future work.

The state-of-the-art models for disease-related information extraction from the literature in the BioCreative V task are also based on CRF [18, 19, 23, 34, 35]. Commonly used features include words, part-of-speech tags, word shape features, syntactic relations, and dictionaries. There was also a report [35] showing that RNNs in the BioCreative V task achieved lower results than CRF. Li et al. [20] used RNNs on the BioCreative II GM corpus to extract gene mentions from abstracts. Jagannatha and Yu [12] applied RNNs to extract entities of disease and medication types in EHRs. In addition to studying diseases, a lot of attention in recent years turned to the problem of mining ADRs from social media. One of the first studies on this subject [17] analyzed user posts regarding six drugs from a health-related social network. Benton et al. [1] analyzed message boards to detect drug events using dictionaries and co-occurrence statistics. Freifeld et al. [6] employed a dictionary-based approach to detect mentions of ADRs in tweets. Several studies used CRF to extract the ADRs from tweets [27, 33].

Most relevant to our work were studies by Metke-Jimenez and Karimi [24]. They applied dictionary-based methods and CRF to identify ADRs from the CADEC corpus. For CRF features, they used bag of words, letter n-grams, and word shapes (e.g., if the token composed of uppercase letters). The CRF outperformed other methods and achieved F1-measure of 60.2% in exact matching exercise. Our work differs from the aforementioned reports in several ways: (i) we focus on all disease-related entities, not only ADRs, since it could be more valuable for finding potentially novel causal relations among diseases and drugs; (ii) we experiment with not only feature-rich CRF-based approach but also with bidirectional LSTM and GRU; (iii) we xplore different hidden layer sizes of RNNs; (iv) we use word embeddings trained both on social media and the scientific literature; (v) we analyze the results to explore the variation in levels of effects across different patient conditions.

3. Approach

We formulate the disease-related entity extraction as a sequence labeling problem. CRF [16] is one of the state-of-the-art methods that takes a sequence of tokens as an input, calculates the probabilities of the predefined labels and selects the one with the maximum probability. We view an opinion as a sequence of tokens and label these tokens using the BIO (**B**eginning **I**nside **O**utside) tagging scheme. We identify BIO tags at the document level.

3.1. Features

We use the following set of features for CRF:

- Word (w): the lemmatized word itself;
- Part-of-speech tag (pos): the part-of-speech tag of each word;
- Suffix and Prefix (sp): the suffixes and prefixes of each word up to 6 characters in length;
- Context (context): three groups of features (x, pos, dict) of two words backward and two words forward from the current word;
- Word Type (wtype): two binary features that indicate whether a current word is a negation (*no, not* or *'t*) and whether all characters are capitalized;
- Dictionary Look-up (dict): if we a match can be found in the text, we mark the match using the BIO scheme. For each of three dictionaries, the token has 3 binary features: is beginning of matched part, is in "tail" of matched part, is out of matched part;
- Cluster-based representation (b): the vector of each word described below;
- Word embeddings (emb): the real-valued vector of each word described below.

We have made the implementation of CRF available at the github repository[1].

3.2. Dictionaries

We use the following dictionaries:

1. Dictionary of terms from the Unified Medical Language System (UMLS) with six disease-related types (333,905 entries);
2. Manually validated dictionary of terms (D_{terms}) from UMLS with semantic types "Finding" and "Mental Process" (6,608 entries);

[1] https://github.com/dartrevan/ChemTextMining

3. ADR lexicon adopted from [27] (13,676 entries);
4. Manually created dictionary of multiword expressions (D_{MWE}) (943 entries);
5. Drug names with synonyms from the Drugbank database[2] (57,879 entries).

UMLS is a repository of biomedical vocabularies developed by the US National Library of Medicine. We have used the 2016AA edition of UMLS[1]. We extracted 562,919 medical terms with synonyms from UMLS with the following semantic types: “Disease Or Syndrome”, “Neoplastic Process”, “Sign Or Symptom”, “Congenital Abnormality”, “Mental or Behavioral Dysfunction”, and “Anatomical Abnormality”. We filtered out entries that were non-English terms, stop-words or body parts. The manually created dictionary contains MWEs starting with *feel*, *able* or *ability* such as “feel tired”, “able to relax”, “ability to move”. In addition to medical terminology, UMLS contains Consumer Health Vocabulary, where terms have semantic types “Finding” and “Mental Process”. However, we found many non-relevant to diseases terms with these types (e.g., *born in Cuba, parents got divorced*). In order to filter non-relevant terms, we calculated the frequency of each term in the Health Dataset (described below). Then we manually selected terms with high frequencies and combined them with synonyms in our dictionary (e.g., *drop in blood pressure*, *breakthrough bleeding*, *increased body weight*).

3.3. Word representations and Unlabeled Data

We used two types of word representations: (i) cluster-based and (ii) distributed (also called word embeddings). We collected a large number of 2,607,505 unlabeled user comments from six resources (this collection is further referred to as the Health Dataset) to induce the word representations. The resources included webmd.com[3], askapatient.com[4], patient.info[5], dailystrength.org[6], drugs.com[7]; we also employed health product reviews from freely available Amazon dataset[8]. Duplicate texts were removed. Each comment was processed with the tokenizer and lowercased.

We used the Brown hierarchical clustering algorithm [2]. This algorithm partitioned all words into a set of 150 clusters[9]. Word embedding models represent each word using a single real-valued vector. Such representation groups together words that are semantically and syntactically similar [25]. We used word2vec from Gensim library[10] to train embeddings on the Health Dataset. We applied Continuous Bag

---

2 https://www.drugbank.ca/

3 http://www.webmd.com

4 http://www.askapatient.com

5 https://www.drugs.com

6 https://dailystrength.org

7 http://patient.info

8 http://jmcauley.ucsd.edu/data/amazon

9 https://github.com/percyliang/brown-cluster

10 https://radimrehurek.com/gensim/

of Words model with the following parameters: vector size of 200, the length of local context of 10, negative sampling of 5, vocabulary cutoff of 10. Below, we refer to our pre-trained vectors as HealthVec (93,526 terms). We also experimented with another published word vector PubMedVec (2,351,706 terms) trained on biomedical literature indexed in PubMed [28].

4. Evaluation and Experiments

In this section, we conduct experiments to demonstrate the effectiveness of our CRF-based approach. We first describe the experimental settings and baselines. We compare CRF to the baseline methods and analyze the effect of different features.

4.1. Experimental Settings

The implementation of CRF was based on the sklearn-crfsuite library[11]. We used WordNetLemmatizer and maximum entropy tagger from the nltk library[12]. Passive aggressive algorithm was used for updating feature weights to train CRF.

Dataset. We use the CADEC corpus [14] that annotated with Drug and Disease entities at the sentence level (1250 posts, 7,632 sentences, 101,486 words). The corpus consists of four predefined disease-related types: ADR (6,318 entities), Disease (283 entities), Symptom (275 entities), Findings (435 entities). Since entities of each type are highly imbalanced in the corpus, we join them into one Disease type. We also reduced entities that fully contained in a larger entity of the same type. The total numbers of Drug and Disease entities were 1,799 and 6,752, respectively. To evaluate our method, we employed 5-fold cross-validation.

Baseline Methods. We evaluated our model by comparing with two baseline methods:

1. A knowledge-based approach that relies on the use of the described dictionaries and based on the exact lookup.
2. Bidirectional RNNs, in particular, LSTM and GRU [4, 9].

Our implementation of the knowledge-based approach is based on the Apache UIMA Ruta[13]. In order to implement RNNs, we used the Keras library[14]. The architecture of our networks and parameters are similar to [12]. We used a standard LSTM or GRU with the tanh activation function on top of the embedding layer. The embedding layer is based on pre-trained word embeddings. Bidirectional RNN has two independent forward and backward chains and the output layer that combines them. We use 100-dimensional hidden layer for each RNN chain. Finally, the combination of RNN chains' outputs is fed into a fully connected layer with softmax activation.

---

[11] https://sklearn-crfsuite.readthedocs.io

[12] http://www.nltk.org/

[13] https://uima.apache.org/ruta.html

[14] https://keras.io/

This layer computes probabilities for each of the Drug and Disease labels and the Outside label. In order to prevent neural networks from overfitting, dropout of 0.5 is used to manage the inputs and the softmax layer. We use categorical cross entropy as the objective function. The batch size is 32. We use Adam [15] with a learning rate of 0.01 and a gradient clipping of 5.0 to optimize the cost of our network. We use a maximum of 70 epochs to train each network.

4.2. Experiments

At the pre-processing step, we performed spell correction[15]. We computed recall (R), precision (P) and F_1-measure (F) in two variants: (i) exact matching following CoNLL evaluation [32] and (ii) partial matching described in [22]. We use both Drug and Disease entities and do not present results of extraction of drugs since CRF and RNN extracts 92% of annotations correctly and the NER problem simply does not present itself. The results of different methods and ablation experiments are shown in Table 1 and Table 2, respectively.

Table 1: 5-fold cross-validation of the proposed methods

	Exact			Partial		
Method	**P**	**R**	**F**	**P**	**R**	**F**
Dictionary-based approach	.503	.502	.494	.836	.546	.625
1-layer GRU, HealthVec	.661	.516	.579	.786	.820	.780
2-layer LSTM, HealthVec	.712	.617	.661	.802	.863	.809
2-layer GRU, uniformly distributed rand. embeddings	.554	.489	.519	.740	.712	.694
2-layer GRU, PubmedVec	.669	.614	.640	.818	.800	.783
2-layer GRU, HealthVec	.719	.619	.665	.795	.871	.809
3-layer LSTM, HealthVec	.718	.629	.670	.801	.872	.812
3-layer GRU, HealthVec	.735	.629	.678	.793	.876	.811
CRF, all features + HealthVec	.702	.680	.691	.852	.790	.794

Table 2: 5-fold cross-validation of CRF with different feature groups

	Exact			Partial		
Method	**P**	**R**	**F**	**P**	**R**	**F**
features: w, sp, pos, context, wtype, b, dict, HealthVec	.702	.680	.691	.852	.790	.794
features: w, sp, pos, context, b, dict, HealthVec	.701	.681	.690	.853	.789	.794

[15] http://norvig.com/spell-correct.html

Method	Exact			Partial		
	P	R	F	P	R	F
features: w, sp, pos, context, b, dict, PubMedVec	.667	.682	.674	.829	.815	.799
features: w, sp, pos, context, dict, b	.667	.677	.672	.828	.812	.796
features: w, sp, pos, context, dict	.664	.672	.668	.828	.812	.797
features: w, sp, pos, context, dict (w/o D_{terms} and D_{MWE})	.665	.667	.666	.829	.804	.793
features: w, sp, pos, context	.651	.631	.641	.817	.778	.772
features: w, sp, pos	.615	.601	.608	.810	.771	.764

The results in Table 1 lead to several observations. First, 3-layer GRUs provide the best results as compared to other networks. Second, CRF achieved the best results in the exact matching exercise over GRU due to CRF's capability of predicting a valid sequence of the output labels. Third, F scores of CRF increased from 69.1% to 79.4% in the partial matching as compared to exact exercise since boundaries of opinion expressions are hard to define. Finally, we investigated the effectiveness of CRF's features. The dictionaries along with vectors HealthVec based on in-domain texts led to the most gain in performance of CRF.

5. Analysis of disease-related entities associated with distinct conditions

Although medical terminology is limited, there are a large number of language expressions to describe conditions. To illustrate the variety of phrases which patients use to describe symptoms or drug reactions, we present a comparative analysis of extracted expressions for seven health conditions. For our analysis, we used 143,244 reviews from drugs.com, where each review corresponds to a drug and a condition for treatment. The number of conditions was 558. CRF extracted 684,567 entities from texts. Then we excluded unigrams and phrases that associated with more than one condition. To discuss subjective feelings of illness or drug reactions, we manually selected MWEs that contain words “feel”, “felt” or “feeling”. We present some examples in Table 3.

Table 3. Examples of MWEs associated with medical conditions.

Condition	Multiword Expressions
Fibromyalgia	electric feeling, felt some joint tightness in my neck, sunburn feel from my arms and legs, feeling extremely disoriented, feeling like I want to sneeze, felt like worms crawling, feeling flare ups of fibromyalgia symptoms, feeling of nails being driven through my feet away, feel groggy and drop off to sleep

Condition	Multiword Expressions
Birth control	felt like I was constantly getting stabbed, feel like I was gonna pass out from the pain, felt like I was being ripped open inside, feel like having a mini surgery for birth control, felt like my head was going to explode, feel like a typical bloated boat walking around, feel like someone has just died, feeling like I was going insane, feels like my body cannot handle additional hormones in my body
Weight loss	feel "in the mood" to eat anything, no longer feel prisoner to the world of sweets, feeling like I was intoxicated, feel drained sun up to sun down, feel overweight, feel starved, feel the effects of reduced appetite and cravings, not feel out of control eating, feel like my appetite is suppressed, feeling of a decreased appetite
Anxiety	feeling like a deer in the headlights, feel like the room was spinning, feel like i am losing my grip, felt a tremendous sense of fear, feel incredibly awkward in social situations, feeling high or disoriented or mentally clouded, feel slightly less coordinated, feel like the constant dialogue in my head, feel like my heart is beating harder, heart feels like it's going to fly out of my chest, feel despair
Panic disorder	feel like my emotions are so flat, feel like i am a weak person, feelings of guilt, fog-like feeling, feeling of lethargic/less energy, feel my heart starting to race, feel depressed/stress/panic/anxiety, feel hot from the inside of my body, feeling overall hopeless, feel gloomy, choking throat feeling, feeling like I was suffocating
Bipolar disorder	feel "zombie-ish", feel irritable or obsessed over things, feel very angry or depressed, feel genuine happiness, feel positive and optimistic, feel over whelmed with grief, feel manic or depressive, feels like something foreign in your body, feels like I have bugs crawling inside my brain, feels like everything kind of moves
Depression	feel a little emotionally numb, feels like an electrical impulse going through my head, feel sad or bothered by little things, feeling "revved" at night, feeling of pressure behind the eyes, feeling like I wanted to scream, feel like my head is sometimes floating above my body, feeling of "disconnection" from my emotions

Several observations can be made based on results in Table 3. First, we observe differences in expressions associated with mental health disorders, i.e. anxiety, panic disorder, bipolar disorder, depression. Panic disorder affects the emotional health of a patient (e.g., *emotions are so flat, hopeless, emotional numbness*), while people with anxiety experience emotional symptoms related to feelings of fear (e.g, *felt scared, fearful thoughts, distracted by irrational fears*). The authors with bipolar disorder experience euphoric mood (e.g., *feel positive and optimistic, feel genuine happiness*), while depressed people feel withdrawn from socializing and hobbies (e.g., *loss of interest in everything*). These examples demonstrate that social media posts contain variable information for NER. Second, women that take birth control pills describe ADRs such as abnormal pain (e.g., *constantly getting stabbed, being ripped open inside, gonna pass out from the pain*) more emotional than patients with fibromyalgia, where muscle pain and muscle spasms are symptoms of the disease (e.g., *muscles feel extremely "tight", worms crawling*). Third, patients very often don't know what they are troubled by and use creative writing. For example, "prisoner to the world of sweets" is used to rephrase the term "sweet craving", "a deer in the headlights" describes a feeling of being frozen, "feel like someone has just died" is used to describe depression, and "constant dialogue in my head" refers to a cognitive process such as rumination. Therefore, there is a need to create domain-specific dictionaries and map informal expressions to medical terms. Finally, there are shared problems for all disorders, e.g. most common ADRs like allergic reactions and rash (e.g., *sunburn*), drug abuse (e.g, *intoxicated*), or lack of control (e.g., *out of control eating, "disconnection" from my emotions, prisoner in your body*).

Our analysis shows that existing resources can integrate MWEs from social media posts to increase understanding of experiences of personalized expressive and explorative writings by patients and create valuable resources for supervised methods using these unique insights.

6. Conclusion

In this paper, we have explored the task of recognizing opinion expressions in social media associated with diseases and drugs. We complied and harmonized user expressions from multiple resources to create a collection we termed the Health Dataset. We used Conditional Random Fields (CRF) and implemented a variety of features based on contextual information, dictionaries, and word representations. We demonstrated the superiority of CRF as compared to a dictionary-based method and recurrent neural networks. We have also demonstrated the variability in emotional level of expressions depending on the type of patient conditions. Our analysis confirmed the need for qualitative methods to interpret informal disease-related expressions and map them onto medical terms. In addition to drug indications and adverse effects, we also plan to annotate beneficial effects, which could lead to the discovery of previously unknown drug effects and new drug repurposing hypotheses. Additional studies are needed to investigate if such effects may be a result of medication usage in combination with other factors such as life style or food. In future studies, we also plan to create and manually annotate a corpus of user reviews about medications, written in Russian. In summary, continuous advancement and improvement in the accuracy of text mining approaches applied to patient reports in social media will have plausible impact in several areas including pharmacovigilance (especially, for new drugs), drug repurposing, and understanding drug effects in the context of other factors such as concurrent use of other drugs, diet, and life style.

Acknowledgments

Work on problem definition, survey of related work, recurrent networks, dictionaries and qualitative analysis of results was carried out by Elena Tutubalina and supported by the Russian Science Foundation grant no. 15-11-10019. Other parts of this work were performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

References

1. *Benton A., Ungar L., Hill S., Hennessy S., Mao J., Chung A., Holmes J. H.* (2011), Identifying potential adverse effects using the web: A new approach to medical hypothesis generation, Journal of biomedical informatics, Vol. 44(6), pp. 989–996.
2. *Brown P. F., Desouza P. V., Mercer R. L., Pietra V. J. D., Lai J. C.* (1992), Class-based n-gram models of natural language, Computational linguistics, 18(4), pp. 467–479.
3. *Chernyshevich M.* (2014), IHS R&D Belarus: Cross-domain extraction of product features using conditional random fields, Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pp. 309–313.
4. *Cho K., Van Merriënboer B., Bahdanau D., Bengio Y.* (2014), On the properties of neural machine translation: Encoder-decoder approaches, arXiv preprint arXiv:1409.1259.
5. *Choi Y., Cardie C.* (2010), Hierarchical sequential learning for extracting opinions and their attributes, Proceedings of the ACL 2010 conference short papers, pp. 269–274.
6. *Freifeld C. C., Brownstein J. S., Menone C. M., Bao W., Filice R., Kass-Hout T., Dasgupta N.* (2014), Digital drug safety surveillance: monitoring pharmaceutical products in twitter, Drug safety, 37(5), pp. 343–350.
7. *Deftereos S. N., Andronis C., Friedla E. J., Persidis A., Persidis A.* (2011), Drug repurposing and adverse event prediction using high-throughput literature analysis, Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 3(3), pp. 323–334.
8. *Gareev R., Tkachenko M., Solovyev V., Simanovsky A., Ivanov V.* (2013), Introducing baselines for Russian named entity recognition, Proceedings of the International Conference on Intelligent Text Processing and Computational Linguistics, pp. 329–342.
9. *Hochreiter S., & Schmidhuber, J.* (1997), Long short-term memory. Neural computation, Vol. 9(8), pp. 1735–1780.
10. *Huang C. C., Lu Z.* (2016), Community challenges in biomedical text mining over 10 years: success, failure and the future, Briefings in bioinformatics, 17(1), pp. 132–144.
11. *Irsoy O., Cardie C.* (2014), Opinion Mining with Deep Recurrent Neural Networks. EMNLP, pp. 720–728.
12. *Jagannatha A. N., Yu H.* (2016). Bidirectional RNN for Medical Event Detection in Electronic Health Records, Proceedings of NAACL-HLT, pp. 473–482.
13. *Jakob N., Gurevych I.* (2010), Extracting opinion targets in a single-and cross-domain setting with conditional random fields. Proceedings of the 2010 conference on empirical methods in natural language processing, pp. 1035–1045.
14. *Karimi S., Metke-Jimenez A., Kemp M., Wang C.* (2015), Cadec: A corpus of adverse drug event annotations, Journal of biomedical informatics, Vol. 55, pp. 73–81.
15. *Kinga D., Adam J. B.* (2015), A method for stochastic optimization, International Conference on Learning Representations (ICLR).
16. *Lafferty J., McCallum A., Pereira F.* (2001), Conditional random fields: Probabilistic models for segmenting and labeling sequence data, Proceedings of the eighteenth international conference on machine learning, ICML, Vol. 1, pp. 282–289.

17. *Leaman R., Wojtulewicz L., Sullivan R., Skariah A., Yang J., Gonzalez G.* (2010), Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks, Proceedings of the 2010 workshop on biomedical natural language processing, pp. 117–125.
18. *Lee H. C., Hsu Y. Y., Kao H. Y.* (2015), An enhanced CRF-based system for disease name entity recognition and normalization on BioCreative V DNER Task, Proceedings of the Fifth BioCreative Challenge Evaluation Workshop, pp. 226–233.
19. *Li D., Afzal N., Mojarad M. R., Elayavilli R. K., Liu S., Wang Y., Liu H.* (2015), Resolution of chemical disease relations with diverse features and rules, Proceedings of the fifth BioCreative challenge evaluation workshop, pp. 280–285.
20. *Li L., Jin L., Huang D.* (2015), Exploring recurrent neural networks to detect named entities from biomedical text, Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data, pp. 279–290.
21. *Liu P., Joty S. R., Meng H. M.* (2015), Fine-grained Opinion Mining with Recurrent Neural Networks and Word Embeddings, Proceedings of EMNLP, pp. 1433–1443.
22. *Loukachevitch N., Blinov P., Kotelnikov E., Rubtsova Y., Ivanov V., Tutubalina E.* (2015), SentiRuEval: testing object-oriented sentiment analysis systems in Russian, Proceedings of International Conference Dialog, Vol. 2, pp. 3–13.
23. *Lu Y., Ji D., Yao X., Wei X., Liang X.* (2015), CHEMDNER system with mixed conditional random fields and multi-scale word clustering, Journal of cheminformatics, Vol. 7(1).
24. *Metke-Jimenez A., Karimi S.* (2015), Concept extraction to identify adverse drug reactions in medical forums: A comparison of algorithms, arXiv preprint arXiv:1504.06936.
25. *Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J.* (2013). Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems, pp. 3111–3119.
26. *Moghaddam S., Ester M.* (2011), ILDA: interdependent LDA model for learning latent aspects and their ratings from online product reviews, Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, pp. 665–674.
27. *Nikfarjam A., Sarker A., O'Connor K., Ginn R., Gonzalez G.* (2015), Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features, Journal of the American Medical Informatics Association, pp. 1–11.
28. *Pyysalo S., Ginter F., Moen H., Salakoski T., Ananiadou S.* (2013), Distributional semantics resources for biomedical text processing, Proceedings of Languages in Biology and Medicine.
29. *Qu L. et al.* (2016) Named Entity Recognition for Novel Types by Transfer Learning //arXiv preprint arXiv:1610.09914.
30. *Stanovsky G., Gruhl D., Mendes P. N.* (2017) Recognizing Mentions of Adverse Drug Reaction in Social Media Using Knowledge-Infused Recurrent Models.
31. *Titov I., McDonald R. T.* (2008), A Joint Model of Text and Aspect Ratings for Sentiment Summarization, ACL, Vol. 8, pp. 308–316.

32. *Tjong K., Sang E. F., De Meulder F.* (2003), Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition, Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003, Vol. 4, pp. 142–147.
33. *Wang W.* (2016), Mining adverse drug reaction mentions in twitter with word embeddings. Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing.
34. *Wei C. H., Peng Y., Leaman R., Davis A. P., Mattingly C. J., Li J., Lu Z.* (2015), Overview of the BioCreative V chemical disease relation (CDR) task, Proceedings of the fifth BioCreative challenge evaluation workshop, pp. 154–166.
35. *Wei Q., Chen T., Xu R., He Y., Gui L.* (2016), Disease named entity recognition by combining conditional random fields and bidirectional recurrent neural networks, Journal of Biological Databases and Curation.
36. *Wong T. L., Bing L., Lam W.* (2011), Normalizing web product attributes and discovering domain ontology with minimal effort, Proceedings of the fourth ACM international conference on Web search and data mining, pp. 805–814.