End-to-End Deep Framework for Disease Named Entity Recognition Using Social Media Data

Zulfat Miftahutdinov
Kazan (Volga Region) Federal University
Kazan, Russian Federation
Email: zulfatmi@gmail.com

Elena Tutubalina
Kazan (Volga Region) Federal University
Kazan, Russian Federation
Email: ELVTutubalina@kpfu.ru

***Abstract*—A growing interest in the natural language processing methods applied to healthcare applications has been observed in the recent years. In particular, new drug pharmacological properties can be derived patient observations shared in social media forums. Developing approaches designed to automatically retrieve this information is of no low interest for personalized medicine and wide-scale drug tests. The full potential of the effective exploitation of both textual data and published biological data for drug research often goes untapped mostly because of the lack of tools and focused methodologies to curate and integrate the data and transform it into new, experimentally testable hypotheses.**

Deep learning architectures have shown promising results for a wide range of tasks. In this work, we propose to address a challenging problem by applying modern deep neural networks for disease named entity recognition. An essential step for this task is recognition of disease mentions and medical concept normalization, which is highly difficult with simple string matching approaches. We cast the task as an end-to-end problem, solved using two architectures based on recurrent neural networks and pretrained word embeddings. We show that it is possible to assess the practicability of using social media data to extract representative medical concepts for pharmacovigilance or drug repurposing.

***Index Terms*—Medical systems, healthcare, deep learning, recurrent neural networks, disease named entity extraction, disease named entity normalization**

I. Introduction

One of the most important and often underused characteristics of modern science and technology is the unprecedented growth in the volume of accumulated data. This observation is particularly true in drug discovery and pharmacotherapy optimization where data has grown due to the proliferation of high throughput methods in chemical synthesis, biological screening, and rapid accumulation of health-related records. The rapidly growing field of *pharmacovigilance* is concerned with healthcare-relevant information that can be collected automatically from publicly available sources.

Patients widely publish messages associated with health information online in social media, discussion groups, and message boards. Such kind of data sources contains a huge amount of information as unstructured textual data, which in one form or another can be related to health conditions. This implicitly indicates status about the user's personal health and his attitude. Thus, information contained in these data sources could be treated as essential for health applications ranging from understanding about patient's mental health to detection of adverse drug reactions (ADRs). The core task of the pharmacovigilance research is the automation of monitoring of various sources in order to identify potential adverse drug events and interactions. Considering that it is highly important to examine the relationship between social media post and health-related factors including user profile or health conditions [1].

Extraction of health-related entities has a potential to increase knowledge acquirement of drug discovery and safety surveillance. In this work, we introduce an end-to-end framework that aims to extract and normalize disease related mentions from user comments in social media. Our framework consist of two components, both based on the deep neural architectures: (i) a model for named entity recognition which is the task of identifying one-word or multi-word expressions in a free-form text that refer to certain classes of interest (i.e., diseases), (ii) a sequence-to-sequence network to generate formal medical terminology conditioned on the input text from social media language. In this paper, we present the general description of our framework and discuss the potential of extracting medical concepts from social media with several case studies. We also demonstrate that deep learning has brought new advances in this field.

The paper is organized as follows. In Section II, we survey related work about natural language processing in the biomedical or health domain. In Section III, we present our deep framework. We demonstrate the extraction results from a collection of user comments in Section IV and conclude the paper with Section V.

II. Related Work

This section gives an overview of the existing text mining and Natural Language Processing (NLP) techniques applications in biomedical and healthcare tasks. In [1], Kotov has considered recent works on social media analysis that shows the opportunity of mining valuable information for healthcare from social media resources.

Main parts of a knowledge acquisition system are automated methods for information extraction from textual data [2]. Thus, most research efforts were concentrated on identification of medical entities such as adverse drug events and health conditions, entity-entity relations, medical concept normalization

[3], [4], [5], [6] and sentiment analysis [7], [8], [9], [10], [11], [12], [13], [14], [15].

Traditionally, many published studies used dictionary-based approaches [16], [17], [18], [19] or rule-based approaches for extraction of adverse reactions [20]. The list of dictionaries includes well-known publicly available resources SNOMED-CT, MedDRA, SIDER, MeSH, Unified Medical Language System (UMLS) and manually created dictionaries of ADR mentions extracted from various resources such as drug labels, records of clinical trials, user reviews on social media. The general limitations of these methods are low recall of information extraction from social media and unavailability to use for under-resourced natural languages.

Most modern studies have applied machine learning techniques for classification of health-related posts or extraction of named entities [21], [22], [17], [22], [4], [6], [23]. For machine learning, the commonly used features are n-grams, part of speech tags, semantic types from UMLS, negation of words, belonging to ADR dictionary, drug names, distributed representations of words. Recently, Huynh et al. [24] proposed a combination of two architectures named Convolutional Recurrent Neural Network (CRNN).

Due to the significant difference in professional medical and social media language models, several studies focused on the problem of normalizing a free-form "description" of a medical concept to its formal representation. Moreover, there is a number of a number of similar tasks described in the literature. In biomedical research, the similar task is to map chemical mentions or genes into MeSH identifiers. In clinical research, the similar task is to map causes of health problems into the codes of the International Classification of Diseases (ICD). To solve these problems, various methods were applied including learning to rank methods [25], dictionaries [26], convolutional neural networks [27], and recurrent neural networks [3].

The sentiment analysis is one of the important subjects of research in the medical field and utilized across a wide range of applications. Sarker et al. [28] used the sentiment analysis for the purpose of evaluating the effectiveness of treatment. Such studies could be helpful for practicing doctors in making decisions about the treatment and participation of the patient in the treatment process could lead to improvement in the quality of medical services provided [29]. In [30], multi-step classification approach was suggested for opinion mining from Patient Opinion service. In [15], [12], authors for the purpose of treatment effectiveness evaluation used classic sentiment analysis of medical records. Along with patient health condition determination sentiment analysis could be applied to reviews about medical goods and drugs [31]. These studies allow researchers to identify the side effects of drugs, and to improve the pharmaceutical companies their medicine production processes. Another application of sentiment analysis is the patient mental health evaluation. In [32], tweets associated with diabetes were analyzed for sentiment determination. In [33], [34], [35], moods of people with cancer were evaluated.

In the recent years, some scientific groups started works on examining the impact of language-derived personality and demographic information to morbidity [36], [13], [14], [37]. As discussed in [38], application of the NLP approaches could benefit in public health research questions. In different studies based on various corpora, strong correlations were found between user profile including linguistic features and official statistics provided by US CDC. In [4], qualitative analysis led to the conclusion that various people describe in different ways their health conditions depending on their demographic. The work of Benton et al. [39] analyzed the applicability of multi-task learning approaches on mental health tasks.

In conclusion, we note that although neural network architectures have become the method of choice for many different applications, many studies or existing systems (such as Open Targets [40]) in biomedical and clinical research still apply baseline computational models and do not consider recent advances in the field of NLP and deep learning.

III. Framework for Disease Named Entity Recognition and Normalization

In this section, we describe the proposed deep neural architectures for the disease named entity recognition (NER) and normalization. As shown in Figure 1, our framework consists of two main neural networks and overall six basic units:

1) The input of our system (marked with a dotted border) is a free-form text.
2) The preprocessing steps (marked with a dashed border) can include tokenization, POS-tagging, lemmatization. The system maps each word (or lemma) occurring in the corpora' dictionary to its vector representation (word embedding).
3) The NER component which based on LSTM-CRF (marked with a solid border). The detailed description is presented in Subsection III-A.
4) The input/output rectangle (marked with a dotted border). The extracted entities are marked in blue.
5) The normalization component uses a sequence-to-sequence network namely encoder-decoder LSTM (marked with a solid border). The detailed description is presented in Subsection III-B.
6) The final output rectangle (marked with a dotted border). The final medical concepts are marked with a blue border.

As is shown in Figure, the first neural network extracts four disease-related entities marked in blue: "left shoulder is almost immobile", "weak feeling", "not sure footed as I walked", "lousy sleeping at night", while the second network maps entities to medical concepts "shoulder stiff", "asthenia", "unsteady when walking", "difficulty sleeping", respectively.

A. Named Entity Recognition

The state-of-the-art machine learning models for NER are Conditional Random Fields (CRF) [41] and deep recurrent neural networks (RNN) [42], in particular, LSTM [43], [44]. Since it is logical to try to unite CRF and RNN, our framework combines these two methods. The intuition behind this it

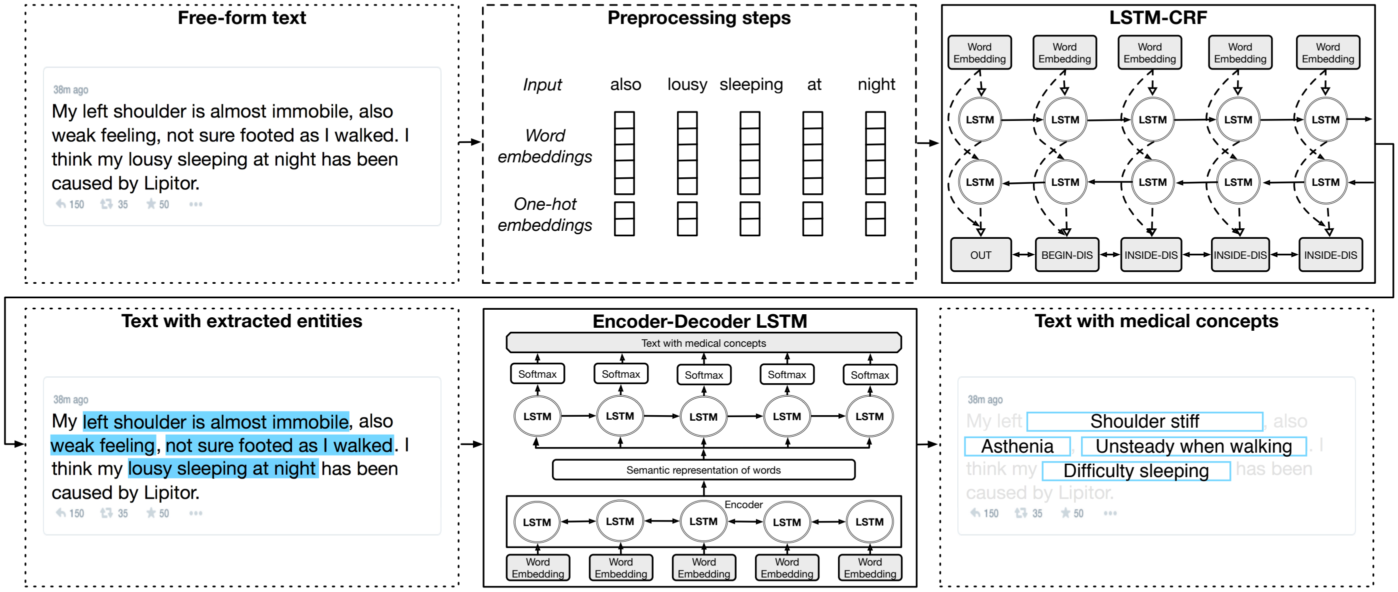

Fig. 1. A graphical representation of our framework

is as follows. First, the distributed representations of words feed into the bidirectional RNN. The representations can be concatenated with one-hot embeddings. The network returns the representation of word's context which feeds into the hidden layer. The number of dimensions equals to a number of tags. The main difference with the basic RNN architecture is that we do not use the output layer with softmax activation function from this layer directly but rather utilize the output of the dense layer for an additional CRF layer to jointly decode the sequence of context tags. In a direct RNN application, especially with LSTM cells, one can get a better model for long input sequences, yet the output layer with softmax will classify every tag independently. CRF can resolve this problem but is less powerful than RNN in modeling the sequence itself. As shown in Figure 1, we determine the boundaries of disease expressions using the BIO scheme, where Begin-DIS, Inside-DIS, and O indicate the beginning, inside and outside entity's tokens.

Our framework utilizes the bidirectional 3-layer LSTM-CRF which was presented in [36] for detecting ADRs from patient reviews about medications. Our model employs existing word embeddings trained on 2.5 millions of user comments from [4].

B. Disease Normalization

For medical text normalization task, also known as terminology association, we utilize an encoder-decoder architecture that successfully used for machine translation applications. In the context of this problem, we translate a text written in a social media language (e.g. "I can't fall asleep all night" and "head spinning a little") to a text written in a formal medical language (e.g."insomnia" and "dizziness", respectively). The encoder-decoder model captures the following intuition: first, we construct a "semantic representation" of a phrase and then "unroll" it in a different language. This is a significant difference compared with traditional NLP studies where unsupervised methods compute semantic similarity measures word by word or apply classification methods to determine whether the free-form text relates to a particular class (i.e., the medical code). The basic architecture of CNN and RNN are successfully applied for classification of tweets into codes of various UMLS resources in English [27].

Our framework utilizes the encoder-decoder model which was firstly proposed in [3] for mapping ICD-10 codes to fragments of death certificates. An important key of the encoder is bidirectional RNNs, where the past and the future context is available at every time step.

IV. Results and Discussion

For training our models, we use a corpus of adverse drug event annotations namely the Cadec corpus [45]. This dataset contains 1,250 posts which contain on average 6 sentences and 81 words. The dataset's annotations contain 7,311 concepts' mentions such adverse effects, symptoms, findings, and diseases which linked to their corresponding concepts in SNOMED and MedDRA. We mark disease-related annotations to one entity type *Disease*. To evaluate our models, we split the Cadec corpus into two datasets, leaving 30% for testing. LSTM-CRF outperformed a feature-rich CRF and achieved the F1-measures of 69.65% and 81.15% on recognition of ADRs in the exact and partial matching exercises, respectively. LSTM achieved the F1-measures of 82.03% on mapping entities to the controlled vocabulary. Please check our published works for details about networks' settings [6], [3], [4].

For qualitative experiments, we use a corpus of 22,513 reviews automatically crawled from three websites: www.drugs.com, www.askapatient.com and www.webmd.com. Each review contains the following fields: a drug name used to treat this disease, and a free-text review, a publication date. The framework automatically extracted 99,188 disease-related entities and linked these expressions to 974 unique medical concepts from SNOMED.

A. Comparison of Social Media Language and Medical Terminology

Examples of social media phrases and their related medical concepts are presented in Table I. Several observations can be made based on examples. First, As shown in this table, simple string matching approaches may not be able to link the social media language "feeling full" to the medical concept "abdominal distension", or link "awake all night" to "cannot sleep at all", since there are no overlapping words. Second, in everyday life patients use many variations in the description of a particular event such as trouble sleeping ("sleep remained an issue", "probs with sleeping", "didn't sleep well"). Finally, within the nominal phrases, we can identify all health-related sentiment adjectives related to health domain ("bad mood", "low mood", "flat mood") reducing false positive and negative errors compared to simple occurrence statistics in large corpora of reviews.

TABLE I
EXAMPLES OF TERMINOLOGY ASSOCIATION.

Extracted entity	Medical concept
feeling full	Abdominal distension
crushing pressure pinning my dreams	Abnormal dreams
feel slightly weaker becoming more brittle	Asthenia
hisssing noise in my ears	Buzzing in ear
awake all night not being able to sleep inability to sleep not sleeping for 5 nights no more sleeping not sleeping 10+ hours	Cannot sleep at all
aggravate mood bad mood low mood flat mood fell into an abyss of depression	Depression mood
struggling to breathe breathing problems impossible to breathe hard to breathe labored breathing	Difficulty breathing
cannot sleep problems sleeping trouble sleeping sleep issue sleep remained an issue didn't sleep well little sleep hard to sleep dont sleep to well probs with sleeping can barely sleep	Difficulty sleeping

B. Using User Comments for Pharmacovigilance

Table II shows top-5 diseases in the comments about Pregabalin, Abilify, and Pradaxa. Pregabalin (Pregabalin) is used to treat pain caused by damage to nerves due to disease or injury. Aripiprazole (Abilify) is used to treat certain mental disorders. Dabigatran (Pradaxa) is used to prevent stroke and harmful blood clots. As shown in this table, most authors experienced ADRs related to weight gain and dizziness.

TABLE II
MOST FREQUENT DISEASE MENTIONS PER DRUG.

Drug name	Most frequent diseases
Pregabalin	Weight gain Dizziness Drowsy Insomnia Difficulty sleeping
Abilify	Depression Weight gain Anxiety Insomnia Lack of energy
Pradaxa	Heartburn Nausea Diarrhea Stomach problem Dizziness

C. Using User Comments for Drug Repurposing

At least part of the side effects of taking medications are often related specifically to the characteristics of the functioning of biological targets. Therefore, the search for new, unexplored biological targets for drug development is an important theoretical as well as practical task. Thus, the development of drugs that attack this biological target will allow the creation of compounds of a new generation (with less toxicity and greater efficacy). We hypothesize that modern text mining technology is capable of extracting reliable assertions between drugs and conditions (diseases) from unstructured data.

Figure 2 depicts the relative frequencies of mentions of the medical concept "weight loss" in crawled reviews about Abilify and Lyrica. Their relative frequencies were calculated as follows. Suppose f_n is the number of reviews about Abilify in the n-th year and wlf_n is the number of reviews about Abilify where a patient mentioned the concept "weight loss" in the same year. Than relative frequency of Abilify and "weight loss" in the n-th year is $\frac{wlf_n}{f_n}$. Relative frequencies for Lyrica were calculated similarly. As shown in the figure, Abilify is stably correlated with weight loss compared to the Lyrica. This analysis of social media posts about Abilify provides implicit evidence for drug association with "weight loss". We note that scientific publications are a reliable source for hypothesis confirmation. In [46], Barak and Aizenberg suggested aripiprazole as a candidate for the treatment of weight gain after manual analysis of literature in 2010.

V. CONCLUSION

In this work, we have tackled a problem often looked in pharmacovigilance studies: can we reliably predict health-related named entities (in our case, diseases and adverse drug reactions) from users' reviews? We have presented the general

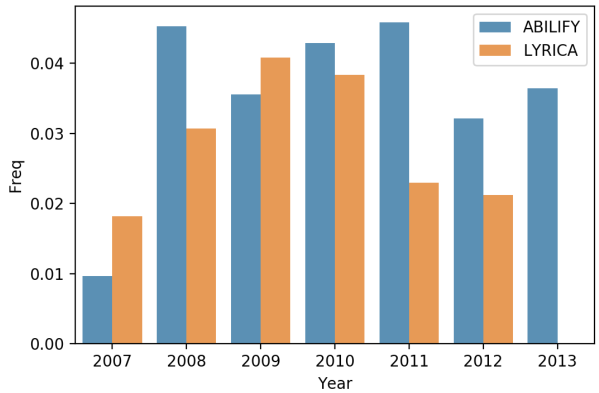

Fig. 2. The relative frequencies of the medical concept "weight loss" in patient reviews about Abilify and Lyrica

framework to this problem, including the joint model LSTM-CRF for extraction of entities and the encoder-decoder LSTM for mapping these extracted entities to corresponding medical terms. Our models are trained end-to-end on a corpus of user reviews with disease-concept annotations. As a result, we have seen that our neural models incorporate and learn extra context information that may lead to interesting observations relevant to the underlying healthcare application. There are several directions for future work. The framework can be used for other similar tasks in biomedical domain (e.g., extraction of chemicals and targets from scientific literature). We would also like to explore more advanced neural architectures.

We believe that in the future, the more healthcare applications will be joined with the predictive power of neural networks, tailored specifically for problems such as chemical-induced disease relation extraction or large-scale phenome-wide association extraction.

ACKNOWLEDGEMENTS

The work on problem definition and neural networks was 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] A. Kotov, "Social media analytics for healthcare," pp. 309–340, 2015. [Online]. Available: http://www.crcnetbase.com/doi/abs/10.1201/b18588-11

[2] V. Solovyev and V. Ivanov, "Knowledge-driven event extraction in russian: corpus-based linguistic resources," *Computational intelligence and neuroscience*, vol. 2016, p. 16, 2016.

[3] Z. Miftakhutdinov and E. Tutubalina, "Kfu at clef ehealth 2017 task 1: Icd-10 coding of english death certificates with recurrent neural networks." CLEF, 2017.

[4] Z. Miftahutdinov, E. Tutubalina, and A. Tropsha, "Identifying disease-related expressions in reviews using conditional random fields," in *Proceedings of International Conference on Computational Linguistics and Intellectual Technologies Dialog*, vol. 1, 2017, pp. 155–167.

[5] S.-P. Choi, S. Lee, H. Jung, and S.-k. Song, "An intensive case study on kernel-based relation extraction," *Multimedia Tools and Applications*, vol. 71, no. 2, Jul 2014. [Online]. Available: https://doi.org/10.1007/s11042-013-1380-5

[6] E. Tutubalina and S. Nikolenko, "Combination of deep recurrent neural networks and conditional random fields for extracting adverse drug reactions from user reviews," *Journal of Healthcare Engineering*, vol. 2017, 2017.

[7] W.-Y. S. Chou, Y. M. Hunt, E. B. Beckjord, R. P. Moser, and B. W. Hesse, "Social media use in the united states: implications for health communication," *Journal of medical Internet research*, vol. 11, no. 4, 2009.

[8] J.-C. Na, W. Y. M. Kyaing, C. S. Khoo, S. Foo, Y.-K. Chang, and Y.-L. Theng, "Sentiment classification of drug reviews using a rule-based linguistic approach," in *International Conference on Asian Digital Libraries*. Springer, 2012, pp. 189–198.

[9] H. Sharif, F. Zaffar, A. Abbasi, and D. Zimbra, "Detecting adverse drug reactions using a sentiment classification framework," 2014.

[10] D. Yalamanchi, "Sideffective-system to mine patient reviews: sentiment analysis," Ph.D. dissertation, Rutgers University-Graduate School-New Brunswick, 2011.

[11] E. Cambria, T. Benson, C. Eckl, and A. Hussain, "Sentic proms: Application of sentic computing to the development of a novel unified framework for measuring health-care quality," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10 533–10 543, 2012.

[12] Y. Deng, M. Stoehr, and K. Denecke, "Retrieving attitudes: Sentiment analysis from clinical narratives." in *MedIR@ SIGIR*, 2014, pp. 12–15.

[13] G. C. M. D. C. Harman, "Quantifying mental health signals in twitter," *ACL 2014*, vol. 51, 2014.

[14] D. Preotiuc-Pietro, J. Eichstaedt, G. Park, M. Sap, L. Smith, V. Tobolsky, H. A. Schwartz, and L. Ungar, "The role of personality, age and gender in tweeting about mental illnesses," in *NAACL HLT*, vol. 2015, 2015, p. 21.

[15] T.-T. Dang and T.-B. Ho, "Mixture of language models utilization in score-based sentiment classification on clinical narratives," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2016, pp. 255–268.

[16] A. Benton, L. Ungar, S. Hill, S. Hennessy, J. Mao, A. Chung, C. E. Leonard, and J. H. Holmes, "Identifying potential adverse effects using the web: A new approach to medical hypothesis generation," *Journal of biomedical informatics*, vol. 44, no. 6, pp. 989–996, 2011.

[17] R. Sloane, O. Osanlou, D. Lewis, D. Bollegala, S. Maskell, and M. Pirmohamed, "Social media and pharmacovigilance: a review of the opportunities and challenges," *British journal of clinical pharmacology*, vol. 80, no. 4, pp. 910–920, 2015.

[18] S. Yeleswarapu, A. Rao, T. Joseph, V. G. Saipradeep, and R. Srinivasan, "A pipeline to extract drug-adverse event pairs from multiple data sources," *BMC medical informatics and decision making*, vol. 14, no. 1, p. 13, 2014.

[19] C. C. Freifeld, J. S. Brownstein, C. M. Menone, W. Bao, R. Filice, T. Kass-Hout, and N. Dasgupta, "Digital drug safety surveillance: monitoring pharmaceutical products in twitter," *Drug safety*, vol. 37, no. 5, pp. 343–350, 2014.

[20] A. Nikfarjam and G. H. Gonzalez, "Pattern mining for extraction of mentions of adverse drug reactions from user comments," in *AMIA Annual Symposium Proceedings*, vol. 2011. American Medical Informatics Association, 2011, p. 1019.

[21] R. Harpaz, A. Callahan, S. Tamang, Y. Low, D. Odgers, S. Finlayson, K. Jung, P. LePendu, and N. H. Shah, "Text mining for adverse drug events: the promise, challenges, and state of the art," *Drug safety*, vol. 37, no. 10, pp. 777–790, 2014.

[22] A. Sarker and G. Gonzalez, "Portable automatic text classification for adverse drug reaction detection via multi-corpus training," *Journal of biomedical informatics*, vol. 53, pp. 196–207, 2015.

[23] A. I. and T. E., "Automated detection of adverse drug reactions from social media posts with machine learning," in *Proceedings of International Conference on Analysis of Images, Social Networks and Texts*, 2017.

[24] T. Huynh, Y. He, A. Willis, and S. Rüger, "Adverse drug reaction classification with deep neural networks," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2016, pp. 877–887.

[25] R. Leaman, C.-H. Wei, and Z. Lu, "tmchem: a high performance approach for chemical named entity recognition and normalization," *Journal of cheminformatics*, vol. 7, no. 1, p. S3, 2015.

[26] Y. Tsuruoka, J. McNaught, J. c. Tsujii, and S. Ananiadou, "Learning string similarity measures for gene/protein name dictionary look-up using logistic regression," *Bioinformatics*, vol. 23, no. 20, pp. 2768–2774, 2007.

[27] N. Limsopatham and N. Collier, "Normalising medical concepts in social media texts by learning semantic representation." in *ACL (1)*, 2016.

[28] A. Sarker, D. Mollá-Aliod, C. Paris *et al.*, "Outcome polarity identification of medical papers," 2011.

[29] A. Coulter and J. Ellins, "The quality enhancing interventions project: patient-focused interventions," *London: The Health Foundation*, 2006.

[30] L. Xia, A. L. Gentile, J. Munro, and J. Iria, "Improving patient opinion mining through multi-step classification." in *TSD*, vol. 5729. Springer, 2009, pp. 70–76.

[31] D. Z. Adams, R. Gruss, and A. S. Abrahams, "Automated discovery of safety and efficacy concerns for joint & muscle pain relief treatments from online reviews," *International Journal of Medical Informatics*, vol. 100, pp. 108–120, 2017.

[32] M. d. P. Salas-Zárate, J. Medina-Moreira, K. Lagos-Ortiz, H. Luna-Aveiga, M. Á. Rodríguez-García, and R. Valencia-García, "Sentiment analysis on tweets about diabetes: An aspect-level approach," *Computational and mathematical methods in medicine*, vol. 2017, 2017.

[33] N. Ofek, C. Caragea, L. Rokach, P. Biyani, P. Mitra, J. Yen, K. Portier, and G. Greer, "Improving sentiment analysis in an online cancer survivor community using dynamic sentiment lexicon," in *Social Intelligence and Technology (SOCIETY), 2013 International Conference on*. IEEE, 2013, pp. 109–113.

[34] P. Biyani, C. Caragea, P. Mitra, C. Zhou, J. Yen, G. E. Greer, and K. Portier, "Co-training over domain-independent and domain-dependent features for sentiment analysis of an online cancer support community," in *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. ACM, 2013, pp. 413–417.

[35] R. G. Rodrigues, R. M. das Dores, C. G. Camilo-Junior, and T. C. Rosa, "Sentihealth-cancer: a sentiment analysis tool to help detecting mood of patients in online social networks," *International journal of medical informatics*, vol. 85, no. 1, pp. 80–95, 2016.

[36] E. Tutubalina and S. Nikolenko, "Automated prediction of demographic information from medical user reviews," in *International Conference on Mining Intelligence and Knowledge Exploration*. Springer, 2016, pp. 174–184.

[37] S. Nikolenko and E. Tutubalina, "Demographic prediction based on user reviews about medications," *Computación y Sistemas*, vol. 21, no. 2, pp. 227–241, 2017.

[38] M. Conway and D. OConnor, "Social media, big data, and mental health: current advances and ethical implications," *Current opinion in psychology*, vol. 9, pp. 77–82, 2016.

[39] A. Benton, M. Mitchell, and D. Hovy, "Multitask learning for mental health conditions with limited social media data," in *Proceedings of the 15th Conference of the EACL*, vol. 1, 2017, pp. 152–162.

[40] G. Koscielny, P. An, D. Carvalho-Silva, J. A. Cham, L. Fumis, R. Gasparyan, S. Hasan, N. Karamanis, M. Maguire, E. Papa *et al.*, "Open targets: a platform for therapeutic target identification and validation," *Nucleic acids research*, vol. 45, no. D1, pp. D985–D994, 2016.

[41] J. Lafferty, A. McCallum, F. Pereira *et al.*, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proceedings of the eighteenth international conference on machine learning, ICML*, vol. 1, 2001, pp. 282–289.

[42] J. L. Elman, "Finding structure in time," *Cognitive science*, vol. 14, no. 2, pp. 179–211, 1990.

[43] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM networks," in *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, vol. 4. IEEE, 2005, pp. 2047–2052.

[44] A. Graves, S. Fernández, and J. Schmidhuber, "Bidirectional LSTM networks for improved phoneme classification and recognition," *Artificial Neural Networks: Formal Models and Their Applications–ICANN 2005*, pp. 753–753, 2005.

[45] S. Karimi, A. Metke-Jimenez, M. Kemp, and C. Wang, "Cadec: A corpus of adverse drug event annotations," *Journal of biomedical informatics*, vol. 55, pp. 73–81, 2015.

[46] Y. Barak and D. Aizenberg, "Switching to aripiprazole as a strategy for weight reduction: a meta-analysis in patients suffering from schizophrenia," *Journal of obesity*, vol. 2011, 2010.