**A CASE STUDY OF EXTRACTION OF ADVERSE DRUG REACTION USING BIDIRECTIONAL LSTM NETWORK FROM UNSTRUCTURED DATA**

Report

Submitted to Indian Institute of Technology(ISM), Dhanbad in partial fulfillment of the requirements for the award of the degree of

Dual Degree(B.Tech+M.Tech)

in

Computer Science & Engineering

by

SUMAN CHATTERJEE

Admission No.:15JE001400

Under the guidance of Dr. Annavarapu Sekhara Rao

Assistant Professor

(Dept. of Computer Science and Engineering)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Indian Institute of Technology (Indian School of Mines), Dhanbad-826004

NOVEMBER 2019

**CERTIFICATE**

This is to certify that  **Mr.Suman Chatterjee** (Admission No. 15JE001400), a student of

M.Sc.Tech./ M.Sc./M.Tech. (Name of Subject/Specialization), Dual Degree (B.Tech+M-Tech )(Computer Science and Engineering), Indian Institute of Technology (Indian School of Mines), Dhanbad has worked under my guidance and completed his Dissertation entitled " **A CASE STUDY OF EXTRACTION OF ADVERSE DRUG REACTION USING BIDIRECTIONAL LSTM NETWORK FROM UNSTRUCTURED DATA** " in partial fulfillment of the requirement for award of degree of Dual Degree (B.Tech+M-Tech )(Computer Science and Engineering) from Indian Institute of Technology (Indian School of Mines), Dhanbad.

This work has not been submitted for any other degree, award, or distinction elsewhere to the best of my knowledge and belief. He is solely responsible for the technical data and information provided in this work.

**(Dr. Annavarapu Chandra Sekhara Rao)**

Assistant Professor and Guide

Department of CSE

Indian Institute of Technology (Indian School of Mines), Dhanbad.

**FORWARDED BY:**

Dr. Haider Banka,

Associate Professor

Head of Department

Department of CSE

Indian Institute of Technology (Indian School of Mines), Dhanbad

**ACKNOWLEDGEMENT**

I would like to express my deep sense of gratitude towards Dr. Annavarapu Chandra Sekhara Rao, Assistant Professor, Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, under his guidance I have learnt many things. He has always been inspiring and has always catered to my doubts and problems with solutions. I thank you sir for your guidance.

I express my gratitude towards all the faculty members for their encouragement, support and blessings.

I would also like to express my gratitude towards my batch mates for their support and well wishes.

Lastly, I would like to thank my family members for their constant support and inspiration and for having enormous faith in me.

Suman Chatterjee Admission No: 15JE001400 Dual Degree (B.Tech+M-Tech )(Computer Science and Engineering) IIT(ISM) Dhanbad

**DECLARATION**

The Dissertation titled **" A CASE STUDY OF EXTRACTION OF ADVERSE DRUG REACTION USING BIDIRECTIONAL LSTM NETWORK FROM UNSTRUCTURED DATA "** is a presentation of my original research work and is not copied or reproduced or imitated from any other person ' s published or unpublished work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions, as may be applicable. Every effort is made to give proper citation to the published/ unpublished work of others, if it is referred to in the Dissertation.

To eliminate the scope of academic misconduct and plagiarism,I declare that I have read and understood the UGC ( Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations, 2018. These Regulations have been notified in the Official Gazette of India on 31st July, 2018.

I confirm that this Dissertation has been checked with the online plagiarism detector tool Turnitin

(http:// / www.turnitin.com) provided by IIT (ISM) Dhanbad and a copy of the summary report/ report, showing Similarities in content and its potential source (if any), generated online through Turnitin is enclosed at the end of the Dissertation. I hereby declare that the Dissertation shows less than 10% similarity as per the report generated by Turnitin and meets the standards as per MHRD/UGC Regulations and rules of the Institute regarding plagiarism.

I further state that no part of the Dissertation and its data will be published without the consent of my guide. I also confirm that this Dissertation work, carried out under the guidance of Dr. Annavarapu Chandra Sekhara Rao, Assistant Professor, Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, has not been previously submitted for assessment for the purpose of award of a Degree either at IIT (ISM) Dhanbad or elsewhere to the best of my knowledge and belief.

Suman Chatterjee Admission No: 15JE001400 Dual Degree (B.Tech+M-Tech )(Computer Science and Engineering) IIT(ISM) Dhanbad

(Forwarded by)

(Dr. Annavarapu Chandra Sekhara Rao)

Assistant Professor

Department of Computer Science and Engineering

**Abstract**

Pharmacovigilance can be defined as the science and experiments referring to the detection, assessment, understanding and prevention of side effects or other drug problems. In hospitals, patients are often subjected to multiple treatments, which can cause adverse effects. AE refers to any negative event occurring at the time a drug is taken, whether it’s identified as a cause of said event or not. If one can establish a relation between the event and the drug, then the relation is said to be an Adverse Drug Event (ADE) or Reaction (ADR).For the aim of identifying ADR mentions, we use medical notes provided in EHR (Electronic Health Records). These notes contain mentions of medical entities like medications, ADR and symptoms. Therefore, A reliable automated extraction system would be of high value.

For this problem, we proposed the use of bidirectional-LSTM’s(Long Short Term Memory) with 1-D CNN layer to classify patient notes at character level and at word level. The 1-D CNN is employed to scale back the training time. In order to improve the performance, we will also feed the network combined word embedding consisting of Pre-trained word2vec 100 dimension word embedding trained on the Twitter ADR Dataset database and character embedding generated by a Char-CNN for Named Entity Recognition. This system requires no handcrafted features or rules, unlike existing systems. The Federal Drug and Administration (FDA) shows the patients that exhibit ADRs have double length of stay and mortality than those of non-ADR patients. Moreover, estimate shows that up to 39% of ADRs in pediatric inpatients are often life-threatening or fatal. A study estimated that between 32% and 69% of drug-related admissions may be preventable. Our system works better than the previous LSTM and CNN models for unstructured data. It has potential that can save more than 39% of the lives lost to drug related events.

Contents

Certificate …………………………………………………………………....i

Acknowledgements ………………………………………………………….ii

Declaration…………………………………………………………………iii

Abstract ……………………………………………………………………iv

1 Introduction ………………………………………………………………..1

1.1 Introduction

1.2 Related work

2 Datasets………………………………………………………….2

3 Architecture ………………………………………………………………….3

3.1 Preprocessing Layer

3.2 Information Extraction

3.3 Evaluation

4 Method……………………………………………………………4

4.1 Bi-LSTM Layer

4.2 Preprocessing layer

4.3 Information Extraction:Combined Word Embeddings

4.4 Classification: 1-D Convolutional Neural Network

4.5 Overfitting

4.6 Hyper-Parameters

4.7 Evaluation Metrics

5 Conclusion ………………………………………………………………….5

6 References …………………………………………………………………..6

***Chapter 1***

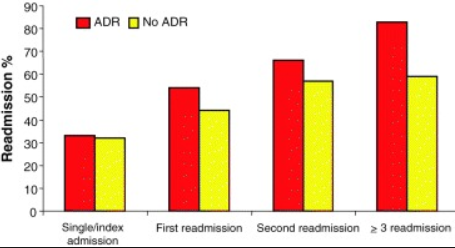
**Introduction**

**1.1Introduction**

Patients are often subjected to multiple treatments, which may cause adverse effects on their bodies. Therefore, it is necessary to see if an Adverse Event (AE) has happened after taking drug/treatment. AE refers to any negative event occurring at the time a drug is taken, whether it’s identified as a cause of said event or not. A negative consequence may deteriorate a patient’s condition or lead to increased healthcare costs. In case one can establish a relation between the AE and the drug, then the relation is considered as an Adverse Drug Event (ADE) or Adverse Drug Reaction (ADR). When multiple drugs are taken at once, the chances of interaction between them increases, where one drug can affect the reaction of others.

**IMPACT:**

* The Federal Drug and Administration (FDA) shows the patients that exhibit ADRs have double length of stay and mortality than those of non-ADR patients. Moreover, estimate shows that up to 39% of ADRs in pediatric inpatients are often life-threatening or fatal. A study estimated that between 32% and 69% of drug-related admissions may be preventable. In the U.S., the cost of ADRs may be up to $136 billion annually.
* Many ADRs are hard to be noticed as they happen to certain types of people in given conditions and they can take a long time to be exposed.
* Healthcare providers conduct clinical trials to discover ADRs before selling the products but normally are limited in numbers. Thus, post-market drug safety monitoring is required to help discover ADRs after the drugs are sold on the market.
* Recently unstructured data such as medical reports or social network data have been used to detect content that contains ADRs. Event and patient reports in the scientific biomedical literature can be found in large volumes and generated rapidly. Social networks are another source of redundant data with unstructured format. While it may contain ADRs that not be clinically useful, a large volume of these data can expose serious or unknown consequences.
* Our system will help in automation of the extraction of the ADRs from drug labels which would increase the efficiency of discovery, which in turn can save millions of lives.



Many doctors have used other resources to spot ADRs, like biomedical journals, social media, and electronic health record (EHR) notes. The ADRs extracted from these resources are an incredible difference to traditional ADR-surveillance systems. However, manually collecting ADRs from these data is too much labor. As such, the utilization of computers may be good option to automatically detect ADRs, but can fail since the data is often unstructured. Therefore, natural language processing (NLP) techniques are utilized for this task.

For the goal of identifying ADE mentions, we use medical notes provided by EHR (Electronic Health Records). These texts contain mentions of medications, ADE (Adverse Drug Event) and symptoms. For the primary step, these terms must be identified and classified within the right category. This classification is called as Named Entity Recognition. During this work we use neural network for NER, using several word representations to enhance the performance. The Second Step is Relation Extraction, which is a model which determines if two entities have a particular relation.

Previous studies utilized traditional ML techniques like condition random field (CRF) or support vector machine (SVM). Currently, deep learning has attracted attention in NLP due to its advantages like better performances and fewer feature engineering compared to other systems. However, only few studies have addressed extraction of ADR- information via deep learning.

## Classification:-

ADRs could also be classified by e.g. cause and severity.

### Cause:

* Type A: Augmented pharmacologic effects - dose dependent and predictable

Type A reaction, which contain mostly 80% of adverse drug reactions, are generally a result of the drug’s primary pharmacological effect (e.g. bleeding when using the anticoagulant [warfarin](https://en.wikipedia.org/wiki/Warfarin)) or a low [therapeutic index](https://en.wikipedia.org/wiki/Therapeutic_index)  (e.g. nausea from [digoxin](https://en.wikipedia.org/wiki/Digoxin)), and they’re therefore predictable. They’re dose-related and typically mild, although they can be serious or even fatal (e.g. intracranial bleeding from warfarin). Such reactions are generally the result of inappropriate dosage, especially when drug elimination is done. The term ‘side effects’ is commonly applied to minor type A reactions.

* [Type B](https://en.wikipedia.org/wiki/Idiosyncratic_drug_reaction): Idiosyncratic

### Severity

The U.S FDA defines a severe adverse event if the patient outcome is one of the following:

* Death
* Life-threatening
* Hospitalization (initial or prolonged)
* Disability - significant, persistent, or permanent change, impairment, damage or disruption in the patient's body function/structure, physical activities or quality of life.
* Congenital abnormality
* Requires intervention to prevent permanent impairment or damage

**ADVERSE DRUG REACTION EXAMPLE TABLE:**

|  |  |  |
| --- | --- | --- |
| **Adverse Drug Reaction** | **Types of Drugs** | **Examples** |
| Anemia (resulting from  decreased production or  increased destruction of red  blood cells) | Certain Antibiotics | Chloramphenicol |
|  | Drugs used to treat malaria  or tuberculosis in people  with G6PD enzyme  deficiency | Chloroquine  Isoniazid  Primaquine |
| Angioedema (swelling of the  lips, tongue, and throat  causing difficulty breathing) | ACE inhibitors | Captopril  Enalapril  Lisinopril |
| Bone fractures | Proton pump inhibitors | Esomeprazole  Lansoprazole  Omeprazole |
| Blood clots | Birth control drugs (all  forms including patches  and pills) | Drospirenone/ethinyl  estradiol  Norelgestromin/ethinyl  Estradiol |
| Confusion and drowsiness | Sedatives | Diphenhydramine |
|  | Antidepressants (especially  in older people) | Amitriptyline  Imipramine |
| Decreased production of  white blood cells, with  increased risk of infection | Certain antipsychotic drugs | Clozapine |
|  | Chemotherapy drugs | Cyclophosphamide  Mercaptopurine  Methotrexate  Vinblastine |
|  | Some drugs used to treat  thyroid disorders | Propylthiouracil |
| Kidney damage | NSAIDs (repeated use of  excessive doses) | Ibuprofen  Naproxen |
|  | Aminoglycoside antibiotics | Gentamicin |
|  | Some chemotherapy drugs | Cisplatin |
|  | Antifungals | Amphotericin B |
| Liver damage | Some analgesics | Acetaminophen |
|  | Some drugs used to treat  tuberculosis | Isoniazid |
|  | Antidepressants | Duloxetine |
|  | Antibiotics | Tetracycline |
| Muscle tissue destruction | **Statins** | Atorvastatin  Simvastatin |
|  |  |  |

In this work, we depend on features learned by neural networks. We use LSTM-based neural network models, not the CNN models employed in a number of the earlier studies. CNN models require pooling on continuous n-grams constructed on a complete sentence so as to get constant length features; here, n is the filter length. This could cause problems for extended length sentences or those that contain relevant clues which are lying relatively far-off from each other.

We utilize Bi-LSTM with 2 different type of pooling techniques to encode variable length features into a constant length feature vector. Theoretically, a Bi-LSTM can store information of past and future words while reading. Therefore, once pooling is applied to the output, we would derive features constituting of information on the complete context (continuous as well as discontinuous) of the whole sentence. In each of those models, we utilize a fully connected 1-D CNN neural network within the output layer.

The two basic characteristics of the models are that each one of them contain one-stage Adverse Drug Events and utilize simple features. None of the chosen features exclusively extracts syntactic information hidden in the very sentence.

**1.2 Related work**

FDA utilizes the Adverse Event Reporting System (FAERS) for detection of adverse events. FEARS includes mandatory reports from pharmaceutical companies and reports that have been submitted to MedWatch directly. The recent approach for FEARS case report review utilizes manually reading the text of the drug labels in so as to confirm whether a patient ADR has been reported earlier or not.

Lipton trained Long Short-Term Memory networks (LSTMs) to classify 128 diagnoses from 13 frequently but erratically sampled clinical results from patients in pediatric ICU. Razavian(2016) used LSTMs to foretell incoming of 133 diseases and conditions simultaneously supported by 18 common lab tests sampled over time. They told that the LSTM learned representations are better than a logistic regression baseline based on hand engineered features. Pham (2017)used LSTMs to model the longitudinal records of diagnoses, medications and procedures and made dynamic predictions of future diagnoses, medications and procedures. They showed enhanced performance over comparative models like SVM.

For unstructured data, Dernoncourt (in 2016)applied bi-directional LSTMs for classifying patient notes. They took two bi-directional LSTM layers, one at character level and also another at word level. Their character level embedding and LSTM purposefully address data sparsity thanks to out-of-vocabulary tokens, misspellings, and different noun forms or verb endings. The 2 layer Bi- LSTMs demonstrated enhanced classification performance, better than state-of-the-art Conditional Random Field (CRF) models.

In recent years, few scientists utilized potential resources from social media to identify ADR. Leaman et al. employed a Lexicon-based approach and utilized 450 comments for Concept/relation extraction system development. Akhtyamova et al. suggested a CNNs model grounded on varied structural parameters. The most vote result in the otuput of the model. Santiso et al. build a deep learning model established on the LSTM to discover ADRs from Electronic Health Records (EHRs).

***Chapter 2***

**DATASETS:-**

* The Twitter ADR Dataset is split into train and test sets.For each set of data of the tweets and annotation are saved in several different files:
* Tweet ID and Annotation files.
* Tweet ID Files:train\_tweet\_ids.tsv,test\_tweet\_ids.tsv.The files have tab separated info of tweet IDs,user IDs and test IDs as demonstrated in the information below.
* 346546465465465 68769687 10238
* The Tweet ID and user ID is utilized for downloading the tweets. The test ID links to the similar annotations in the annotation file.
* Annotation Files: train\_tweet\_annotations.tsv, test\_tweet\_annotations.tsv
* These files contain tab separated info of the details of the annotations including: text ID, start offset, end offset, semantic type, annotated text, related drug and target drug. The next line is an example annotation line. We have to note that the related drug is the (properly spelled) drug that was used as a keyword in Twitter search query and the target drug is the drug that the current annotation (ADR or Indication) is targeting. Target drug can be different from the related drug in cases where there are more than one drug mentions in a tweet.
* 10238 13 34 ADR Restless Leg Syndrome fluoxetine fluoxetine

***Chapter 3***

**Architecture:**

* Preprocessing layer
  + Sentence boundary detector
  + Tokenization
  + Normalization
  + Part of Speech Tagging
  + Shallow Parser
  + Named Entity Recognition
* Information Extraction Layer
  + Feature Extraction Models
    - Combined Word Embedding-Character level and Word Level
  + Classification
    - Bidirectional-LSTM model
    - 1-CNN layer with dropout technique for reducing overfitting.
* Evaluation Metrics and Significance test
* Training & Hyperparameters

***Chapter 4***

***4.1 Bi-LSTM Layer***

The recurrent neural network or RNN is a powerful deep learning technique for modeling sequential data. It has a network of loops, allowing info to persevere throughout the sequence. However, it can suffer from vanishing or exploding gradient problems for events with longer sentences.

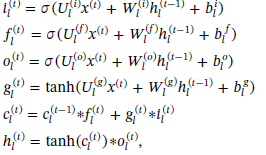
An LSTM network perseveres to overcome this disadvantage by utilizing the gate & memory mechanism. A LSTM layer simply provides another means of computing a hidden state, which introduces a new structure known as a memory cell (c(t)), and three gates: input (i(t)), output (o(t)), and forget ( f (t)) gates. These gates constitutes of a sigmoid activation fn. and are depended on for regulating the memory cell information. The ultimate output of an LSTM layer is calculated supported by the new memory cell states.

Suppose *x1 x2 x3* …..*xm* is the sequence of feature vectors of a sentence,

where *m* is the length of the sentence and is the vector obtained

by concatenating all the feature vectors for the *tth* word. Let *hl(t-1)* and *cl(t-1)* be the previous hidden and cell states of the LSTM (LSTMi), respectively,

then the current hidden state (*htl* ), cell state(*cl(t)* ) and the output of the Bi-LSTM(*z(t)* ) can be computed as follows:

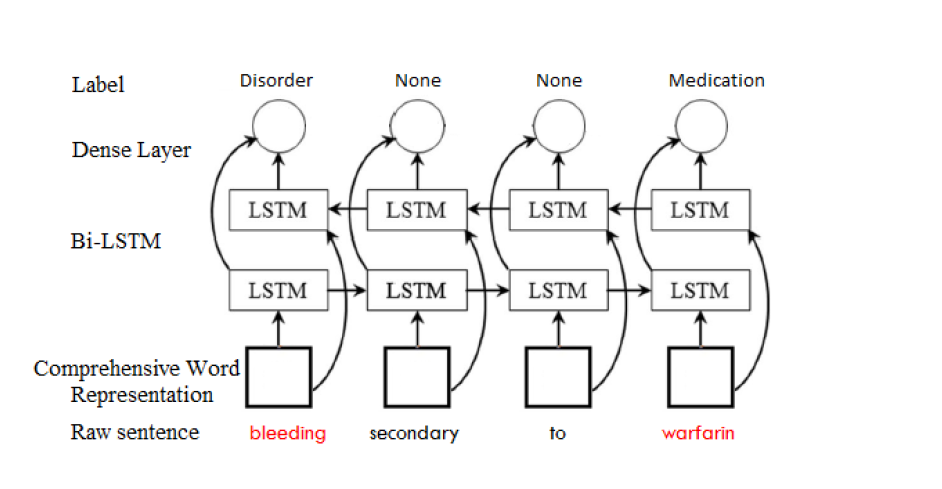


are learning parameters

for LSTM*l*. Here, d is the number of dimensions of an input feature vector and N is the hidden layer size.*hl(t)* is the output of LSTM*l*  at time step t. We compute *hr(t)* in similar manner to *hl(t)* by reversing the words of a sentence. A different LSTM*r* is utilized for this computation. The

final output for the tth word for the Bi-LSTM layer would be:

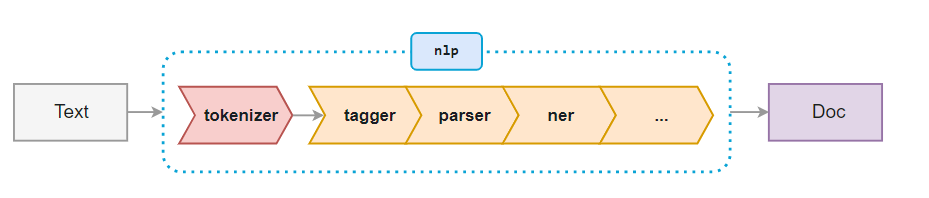


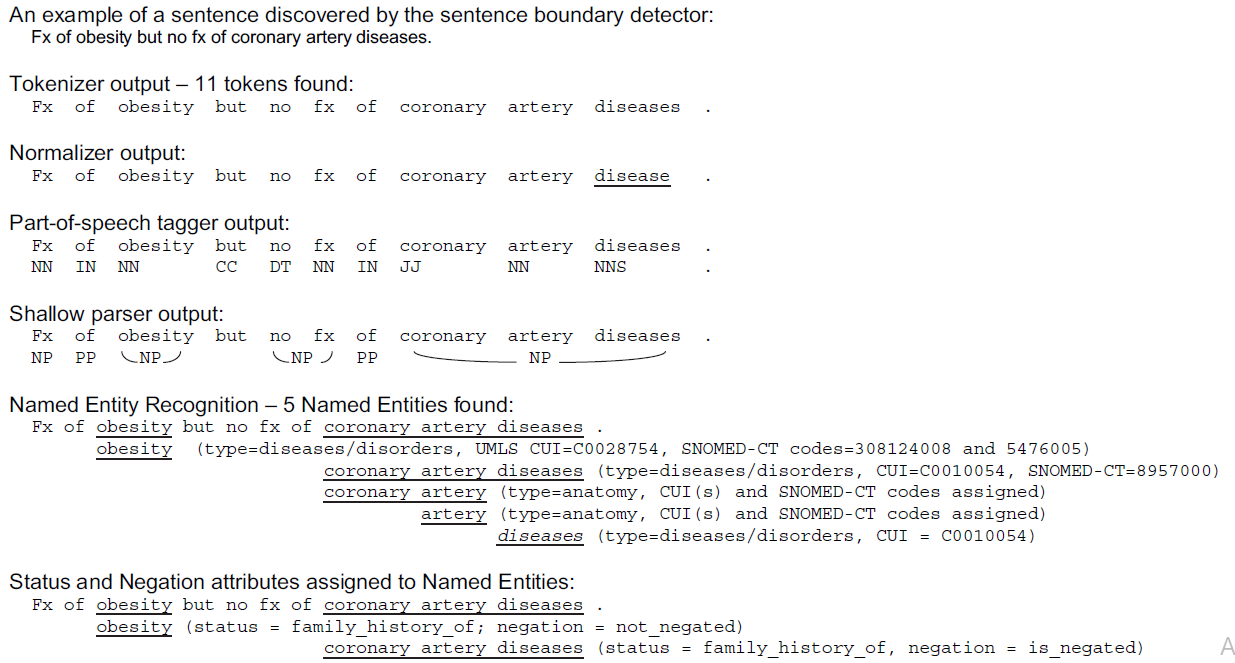


***4.2 Preprocessing Layer***

The following steps are used for preprocessing:

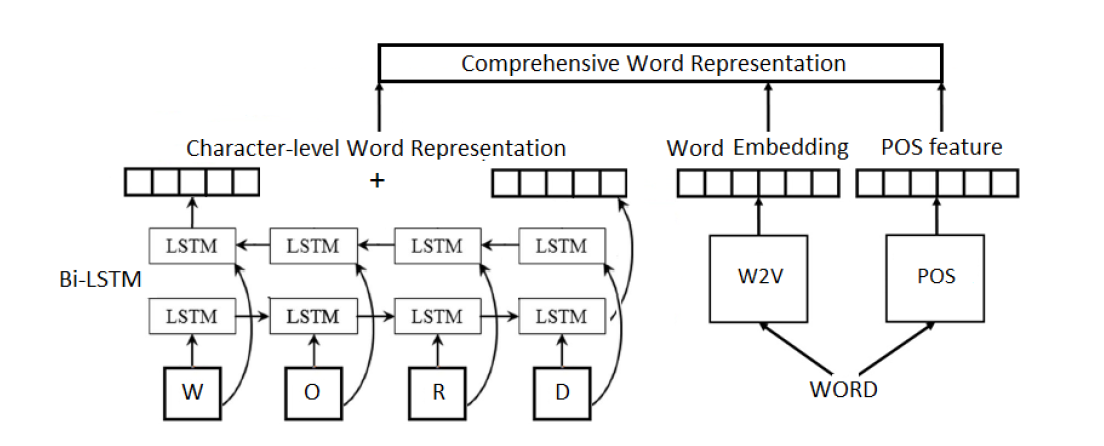
1. We first preprocessed the training dataset in 7 steps: (1) sentence partition, (2) stop words removal, (3) tokenization, (4)Normalization and (5) part-of-speech (POS) tagging.
2. **Sentence Boundary Detector-**It predicts whether a period, question mark, or exclamation mark is the end of a sentence. It is done using the given spacy nlp library .
3. The *Genia tagger* is used for tokenization*.* All digits are normalized and all characters are changed to lowercase.
4. "Tokens" are usually individual words (at least in languages like English) and "tokenization" is taking a text or set of text and breaking it up into its individual words. These tokens are then utilized for the input of other types of analysis or tasks, like parsing (automatically tagging the syntactic relationship between words).
5. **Normalization-**provides a representation for each word in the input text that is normalized with respect to a number of lexical properties. Normalization is used to map multiple mentions of the same word that don’t have the same string representations in the input data.
6. **Part-Of-Speech Tagging-**Assigning word types to tokens, like verb or noun.
7. **Named-Entity Recognition-**Labelling named "real-world" objects, like drugs,diseases,etc..

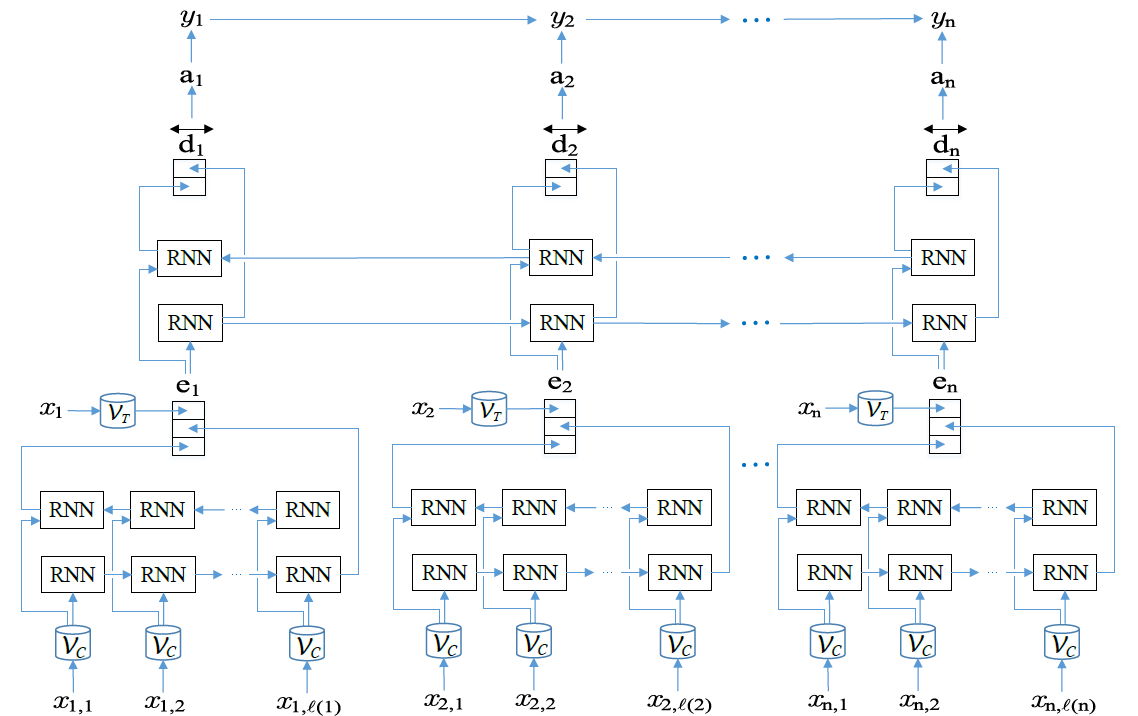
******

******

***4.3 Information Extraction:Combined Word Embeddings***

* For NLP applications, recurrent neural network models are most used together with word embeddings.
* The word embedding is designed to capture semantic similarity of words. The embeddings are significant real-valued vectors of configurable dimension, & semantically similar words generally have close embedding vectors.
* Neural language modeling tools such as word2vec can learn embedding vectors from an unlabeled large text corpus, based on the word’s context in different sentences.
* For word embedding, we experimented with pre-trained word vector on general domain corpus and in-house-trained word vector on clinical notes from MIMIC-III database using word2vec tool.
* The character embedding were generated by a convolutional neural network (CNN) that runs over the characters of a given token.
* The character level embedding and LSTM aim to address data sparsity due to out-of-vocabulary tokens, misspellings and different noun forms or verb endings.

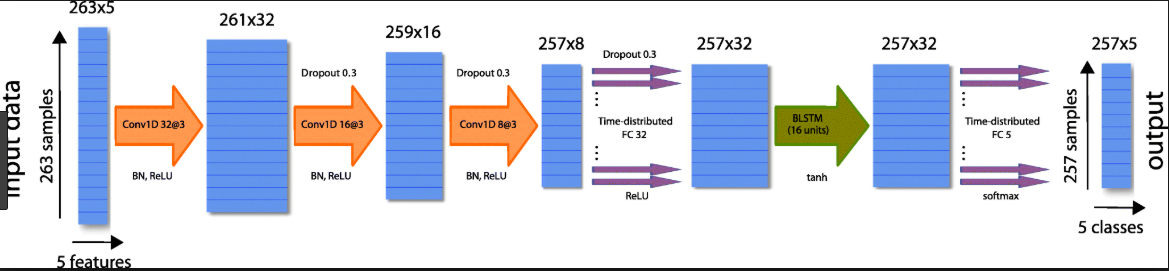


******

***LSTM with word embeddings***

***4.4Classification:1-D Convolutional Neural Network***

* The network starts with a convolutional layer with Rectified Linear Units (RLUs). A RLU accepts an input and returns the initial input if it’s larger than 0, otherwise it returns 0.
* The convolutional filters normally have the identical width because the word vectors, thus, produce feature maps with only 1 column.
* The network is stacked with the help of a max pooling layer that chooses the max element from all columns.
* The last layer is a feedforward layer to an output layer with either sigmoid or softmax activations based on whether the classification is binary or multinomial.



***CNN applied with B-LSTM***

* 1. ***Overfitting:***

For RNNs like LSTMs, overfitting could also be a significant problem. To handle such an issue on sentence LSTM and other LSTM models developed during this study, the dropout technique is utilized to randomly drop the values of a portion (50% in our experiment) of hidden units within the output of the pooling layer during training.

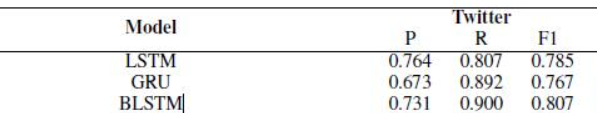
Dropout prevents co-adaptation of those hidden units by sampling from a large number of different ‘‘thinned” networks, thus reducing overfitting and resulting to significant improvements over other regularization methods.

***4.6 Hyper-parameters:***

* Character embedding dimension: 25
* Character-based token embedding LSTM dimension:25
* Word embedding dimension: 100
* Label prediction LSTM dimension: 100
* Dropout probability: 0.5

***4.7 Evaluation Metrics AND Results:***

* To evaluate the efficiency of the 2 models, we calculate precision, recall, and F1-score. Let TP be the number of true positives, FP the number of false positives, and FN the number of false negatives. Precision, recall, and F1-score are defined as follows: **precision = TP/TP+FP=0.731**
* **Recall = TP/TP+FN =0.900,**
* **F1-score = 2\*precision\*recall /precision + recall=0.807.**
* Intuitively, precision is the proportion of the predicted drug labels that are ADR labels, recall is the proportion of the ADR labels that are correctly predicted, and F1-score is the harmonic mean of precision and recall.



***Chapter 5***

***Conclusion***

For this problem, we proposed the use of bidirectional-LSTM’s(Long Short Term Memory) with 1-D CNN layer to classify patient notes at character level and at word level. The 1-D CNN is employed to scale back the training time. In order to improve the performance, we will also feed the network combined word embedding consisting of Pre-trained word2vec 100 dimension word embedding trained on the Twitter ADR Dataset database and character embedding generated by a Char-CNN for Named Entity Recognition. This system requires no handcrafted features or rules, unlike existing systems.

Our system works better than the previous LSTM and CNN models for unstructured data. It has potential that can save more than 39% of the lives lost to drug related events. This study discusses comparison between methods in multiple cases. Analysis of the results demonstrates points such as: imbalance and noise affect all models negatively, the Advice interaction class is the easiest to foretell, repeat mentions of other drug names affect all models negatively, and models can likely to conduct incorrect identification in extended sentences.

***Chapter 6***

**References :-**

[1] Nikfarjam A, et al. Journal of the American Medical Informatics Association 2015;0:1–11. doi:10.1093/jamia/ocu041

[2] Zhang, Y., Chen, Q., Yang, Z. *et al.* BioWordVec, improving biomedical word embeddings with subword information and MeSH. *Sci Data* **6,**52 (2019). https://doi.org/10.1038/s41597-019-0055-0

[3] Adverse Drug Event Detection Using a Weakly Supervised Convolutional Neural Network and Recurrent Neural Network Model Min Zhang 1,2 and Guohua Geng 1,\*.

[4] An attention-based effective neural model for drug-drug interactions extraction Wei Zheng1,2, Hongfei Lin1\*, Ling Luo1, Zhehuan Zhao1, Zhengguang Li1,2, Yijia Zhang1, Zhihao Yang1 and Jian Wang1 Zheng et al. BMC Bioinformatics (2017) 18:445 DOI 10.1186/s12859-017-1855-x

[5] I. Segura Bedmar, P. Martinez, D. Sánchez Cisneros, The 1st DDIExtraction-2011

challenge task: Extraction of Drug-Drug Interactions from biomedical texts, in:

Proceedings of the 1st Challenge Task on Drug-Drug Interaction Extraction 2011,

SEPLN 2011 Workshop, Spain, 2011

[6] I. Segura-Bedmar, P. Martínez, M. Herrero Zazo, SemEval-2013 Task 9: Extraction of Drug-Drug Interactions from Biomedical Texts (DDIExtraction 2013), in: Second Joint Conference on Lexical and Computational Semantics (∗SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013), Association for Computational Linguistics, Atlanta, Georgia, USA, 2013, pp. 341–350.

[7] R. Kavuluru, A. Rios, T. Tran, Extracting drug-drug interactions with word and

character-level recurrent neural networks, 2017 IEEE International Conference on

Healthcare Informatics (ICHI), IEEE, 2017, pp. 5–12.

[8] S. Lim, K. Lee, J. Kang, Drug drug interaction extraction from the literature using a recursive neural network, PLOS ONE 13 (1) (2018) 1–17.

[9] M.F.M. Chowdhury, A. Lavelli, Exploiting the Scope of Negations and

Heterogeneous Features for Relation Extraction: A Case Study for Drug-Drug

Interaction Extraction., in: HLT-NAACL, 2013, pp. 765–771

[10] Zhang, Y.; Yuan, H.; Wang, J.; Zhang, X. YNU-HPCC at EmoInt-2017: Using a CNN-LSTM Model for Sentiment Intensity Prediction. In Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Copenhagen, Denmark, 8 September 2017; pp. 200–204.

[11] Yoon, J.; Kim, H. Multi-Channel Lexicon Integrated CNN-BiLSTM Models for Sentiment Analysis.In Proceedings of the 29th Conference on Computational Linguistics and Speech Processing (ROCLING), Taipei, Taiwan, 27–28 November 2017; pp. 244–253.

[12] Qu, L.; Gemulla, R.; Weikum, G. A Weakly Supervised Model for Sentence-Level Semantic Orientation Analysis with Multiple Experts. In Proceedings of the Joint Conference on Empirical Methods in Natural language Processing and Computational Natural Language Learning. Association for Computational Linguistics, Jeju Island, Korea, 12–14 July 2012; pp. 149–159.