Proceedings of the APAN – Research Workshop 2019
ISBN 978-4-9905448-9-8

Meaningful Information Extraction from Unstructured Clinical Documents

[1]Asim Abbas, [2]Muhammad Afzal, [1]Jamil Hussain and [1]Sungyoung Lee.

***Abstract*— Medical concept and entity extraction from the medical narrative unstructured documents is the crucial step in most of the contemporary health systems. For the extraction of medical concepts and entities, the Unified Medical Language System (UMLS) Metathesaurus is a big source of biomedical and health-related concepts. Recently various tools like Sophia, MetaMap and cTAKES, and many other rules-based methods and algorithm like Quick UMLS etc. have been developed which are performing a successful role in the process of medical concept extraction. The goal of this paper is to design a generic algorithm to identify a package consisting of standard concepts, their semantic types, and entity types on the basis of medical phrases and terms used in the clinical unstructured documents. The proposed algorithm implements the UMLS terminology service (UTS) and customizes to extract concepts for all the meaningful phrases and terms used in the narratives and determine their semantic and entity types in order to find exact categorization of the concepts. The proposed algorithm has produced a very useful set of results to use for labeling the biomedical data, which could in term be used for training data-driven approaches such as machine learning.**

***Index Terms*— UMLS (Unified Medical Language System), Information Extraction, Medical Narratives, Entity Recognition, Semantic Types.**

I. Introduction

ONE of the fundamental and cornerstone stone processes in health informatics system is entity and concept extraction from the medical narrative unstructured documents [1]. Automatic identification of medical concept, entity and semantic relation between is useful for enabling superior and accurate clinical decision support. Currently, distinct tools such as MetaMap [8] and cTakes [9] have been extensively used in research and industry applications for extracting medical concepts.

A knowledge-driven approach is widely used to medical concept identification and extraction build upon various biomedical and health vocabulary such as Unified Medical Language System (UMLS) Metathesaurus [7]. UMLS Metathesaurus is one of the knowledge sources contain 6M names over 100 terminologies, 1.5M concepts, and 8M relations. It is used as prior knowledge source by a different system for concept identification, based on exact match retrieval from lemmatize and original reports. The concept can be used in, among others clinical event detection, information retrieval, question answering and parsing tasks. The vocabulary based system performance relies on three factors (a) The vocabulary size, (b) Used of an algorithm for finding a match, and (c) Presence of large dataset scalability [2].

In this work, we proposed an algorithm that relies on exact matching to terms in UMLS Metathesaurus to extract medical concepts, semantic types, and entity types from unstructured clinical documents. An application programming interfaces (APIs) is provided by UMLS to query the UMLS Metathesaurus data as in Figure.1 for each medical terms. The aim of this task is to introduce the medical text processing to the expansive Natural Language Processing community and to reinforce ongoing research in NLP (Natural Language Processing) methods used in the medical domain. The second purpose of this task is to suggest a generic algorithm for mapping concept, its semantic type and entity type from UMLS to each medical term.

The rest of the paper formulated as follows: an overview of the related work is presented in next section II. The proposed algorithm and framework is described in section III, result and discussion in section IV and the conclusion is drawn in section V.

II. Related Work

In a study [10], the author proposed a rule-based method for automatic concept ex-traction from clinical unstructured documents and mapping these with their semantics type using UMLS Metathesaurus. The proposed method contain two

This research was supported by the Ministry of Science and ICT (MSIT), Korea, under the Information Technology Research Center (ITRC) support programme (IITP-2017-0-01629) supervised by the Institute of Information & Communications Technology Planning & Evaluation (IITP). This work was supported by an IITP grant funded by the Korean government (MSIT) (no.2017-0-00655).

[1] Department of Computer Science and Engineering, Kyung Hee University, Yongin 446-701, South Korea.
*Corresponding Author: Sungyoung Lee
Email Address:{asimabbasturi, jamil, sylee}@oslab.khu.ac.kr
[2] M. Afzal is with Department of Software, Sejong Uninversity, South Korea (e-mail: mafzal@sejong.ac.kr).

sub-module that is text-pre-processing and rules templates. Rules templates further divided into two rules, rule one to identify composite medical concept and rule two for concept mapping with UMLS to identify medical concepts such as problem, treatment, and tests.

In a study [2], the author presented a system an efficient clinical concept extraction in Electronic Medical Record (EMR). An algorithm proposed based on exact matching of clinical terms in reports to determine the longest common word sequence matching, novel prefix indexing and deep parsing based negation detection.

In a study [4], the author developed an approach based on Domain Knowledge and Machine Learning to extract medical concept. For Domain Knowledge, MetaMap tool is used to get important features from text and automatically find terms existent and semantic concepts in UMLS Metathesaurus. In Machine Learning, approach conditional Random Fields (CRF) model is used to detect the clinical name entities. GENIA tool has been used to extract lexical features such as Part Of Speech tags, and Lemmatized version of words.

III. PROPOSED METHODOLOGY

We looked at 2018AB version of UMLS Metathesaurus containing over 1 million concepts. An exact matched term to the UMLS Metathesaurus approach has followed. The proposed algorithm is using Web services API provided by UMLS Terminology Services (UTS) extract single or a list of concept, its Semantic Type and Entity type for each term as shown in the Table1. The big challenge in the algorithm, which we faced, is presenting a value against each specific term, which has many concepts, and then each concept has single or many semantic types. Where semantic type and concept name were repeating for a term. Therefore, for handling this issue we used key value pair logic to hold a list of value for each key. In our algorithm, five main processes are performing Pre-Processing, POS Filtering, Finding Terms, Concept Extraction, Semantic Type Extraction, Entity Type Extraction,

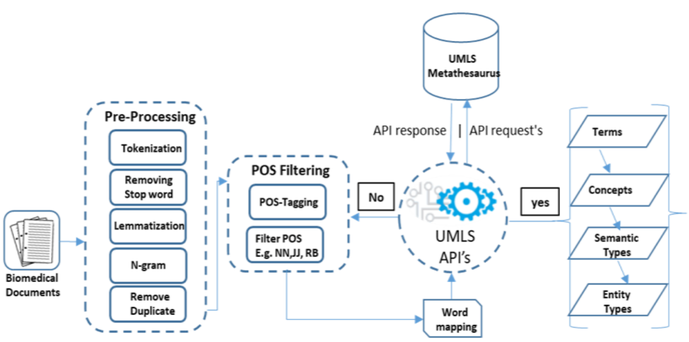

Fig. 1. UMLS Metathesaurus Information Extraction Framework

and Example Scenario of information extraction.

A. Pre-Processing

In pre-processing step, we mine one or more documents D where |D| = {d1, d2, d3…...dn} then we tokenize each document di into sentences then sentences into words |W| where |W|= {w1, w2, w3…....wn}. We apply filtration removing stop words such as "the", "much", "themselves", "this", "from", "on", "off" etc. from the set |W|. These words having no meaning or useless to convey a message. We applied lemmatization to improve the performance on unseen and ambiguous words. Lemmatization is treated more accurate results as compare to stemming. After that, we apply N-gram, which is a set of a co-occurring word in a sentence, for finding n-Gram (see Eq.1).

$$n = x \sim (N - 1) \qquad (1)$$

Where $\sim$ denote the subtraction of a scalar (N-1) from every element of the vector $x = \sum_{k=0}^{n} W_k$ Where denotes the number of words in a sentence. In the last removing of duplicate word process performed. Duplicate words are those word which are repeating in the document due to which document normalize is necessary to avoid ambiguity.

Algorithm.1. UMLS Metathesaurus Information Extraction Algorithm

```
Algorithm : Concept, Semantic Type and Entity Type Extraction

Input: Medication Narrative Document ← {d1, d2, d3 ... dn}  # set of documents di
Output: set of Term, Concept, Semantic Type, Entity Type

1.  wordList ← newArrayList<>
2.  conceptMap ← newMultiMap<term, concept >
3.  semanticTypeMap ← newMultiMap<concept, semanticType >
4.  entityTypeMap ← newMultiMap<semanticType, entityType>
5.  Doc : Read Document
6.  wordList ← Pre-Processing
7.  wordList ← Part of Speech-Filtering
8.  for each term  in wordList, do
        String: cui, concept ← searchConcept (parameter: term)
        conceptMap<k, v> ← term, concept
                cuiList ← cui
    end
9.  for each cui  in cuiList, do
        String: conceptName, semanticType ← searchSemanticType (parameter: cui)
            semanticTypeMap<k, v> ← conceptName, semanticType
                semanticTypeList ← semanticType
    end
10. for each type  in semanticTypeList, do
        String: semanticType, entityType ← searchEntityType (parameter: type)
            entityTypeMap<k, v> ← semanticType, entityType
    end
11. for each term  in wordList, do
        word ← conceptMap(key=term)
        if  word ← NULL, do
            Print (" ")
        end
        else
            Print (term)
            for each concept  in conceptMap.get(term), do
                Print (concept)
                if semanticTypeMap.get(concept).size > 1, do
                    for semanticType in semanticTypeMap.get(concept)
                        Print (semanticType, entityTypeMap(semanticType))
                    end
                end
                else
                    Print ( semanticTypeMap(concept)), entityTypeMap(semanticTypeList.get(0)))
                end
            end
        end
    end
```

B. POS-Filtering

POS filtering module provide support to the pre-processing step in the sense to remove meaningless word from a bag of words received from pre-processing such as "effective", "improved", "rather" etc. In this step, we tagged all the words with part of speech tag. We applied filter using regular expression to choose only noun, adjective and adverb from a bag of words and removed all the verbs as shown in RE.1 and Fig.1. In RE.1. NN means to allow all the nouns, JJ is all the adjectives and RB is the entire adverbs.

Bag of Words = "{<NN.*>}{<JJ.*>}{<RB.*>}"   RE (1)

C. Finding Terms

In the above Eq.1, x represents a list of words. We mapped each x_i to the UMLS Metathesuerus, if x_i found in UMLS the sequence of a term will add to the list Term $|t|$ else x_i will be ignored to the domain expert.

If $x_i \in \text{UMLS}$ then,
Term $|t|$ will add to the list (see Algo 1 Step 8)
Else
x_i Omitted from the list

D. Concepts Extraction

A concept is the meaning of medical terms and each meaning contain different names. Over one million concepts contain Metathesaurus over 100-source vocabulary [11]. Each concept has assigned a unique identifier (CUI), when a new concept is inserted to the Metathesaurus structure. The importance of Metathesaurus is to understand the proposed meaning of each name and to link all the names from all of the source vocabularies, which gives the same meaning called synonyms. In the proposed algorithm for each term t_i single or a list of a concepts is extracted from the UMLS Metathesaurus as showing in Table 1. In the section, 3.2 we have discussed that all the terms $|t|$ after matching with UMLS Metathesaurus will store in list (see Algo 1 Step 8). In Eq.2, C_n is presenting a list of concept available in Metathesaurus for each term $f_1(t_i)$. In the proposed algorithm, we are storing value in a key value format such as <key, value>. We assigned key="term" t_i and value="concept" C_i (see Algo 1 Step 8). Based on each term t_i a single or a list of concepts extracted from array list (see Algo1 Step 11) and Table I.

$$f_1(t_i) = C_n \qquad (2)$$

Where Cn = {c1, c2, c3…. Cn} sequence of concept

E. Semantic Type Extraction

Another knowledge source of UMLS is Semantic Network, contains 135 semantic types (STY's) and 54 semantic relationships for classifying and categorizing of all concepts represented in the UMLS Metathesaurus [1][12]. Each concept C_i contained at least one semantic type (STY) very few concepts assigned five semantic types (see Eq.3). A multifaceted or has inherent ambiguity concept contain more than one STY such as "Febrile Convulsion" is a concept of "Finding" and "Disease or Syndrome" [5]. Metathesaurus concepts give an interpreted and obvious meaning with semantic type and Semantic relation [11]. Semantic Type categorizes the concept in three broad categories such as Medical Problems (pathologic function, cell or molecular dysfunction, disease or syndrome etc.), Medical treatment (therapeutic or preventive procedure, antibiotics, clinical drugs etc.), and Medical Test (laboratory procedure or diagnostic procedure) [4]. In the proposed algorithm, semantic type value is also store in a key value format as discussed in section D. We set key="concept" C_i and value ="semantic types" S_i and map each to the relative concept that is < C_i , S_i > (see Algo1 Step 9). Based on each concept at least one semantic type will extract (see Algo1 Step11) and Table I. In this algorithm, we are focusing on extracting the entire concept, its semantics type and entity Type for a term in unstructured clinical documents.

$$f_2(C_1) = S_n \qquad (3)$$

Where, Sn = {S1, S2, S3…. Sn} sequence of STY's

F. Entity Type Extraction

An entity type is showing the parent relation for the concept. The concept meaning is presented in more standard and obvious as compared to the semantic type. For each concept, it is possible to have more than one semantic type but having only one entity type as in (equ4). In UMLS Metathesaurus entity type assigned the most specific meaning to the concept [11].In the proposed algorithm Entity Type $|E|$ is also list in key value format. We assigned key=" STY" S_1 and value="Entity type" E_i in a list as shown in (see Algo1 Step10). Against each semantic type STY S_1 entity type E_1 extracted as shown in the Table I.

$$f_3(S_1) = E_1 \qquad (4)$$

Where, i = {1, 2, 3…..n}

G. Example Scenario of Information Extraction

In Fig.2. An example scenario has drawn in the form of hierarchal tree that presented how our algorithm extract information from UMLS Metathesaurus for medical terms.

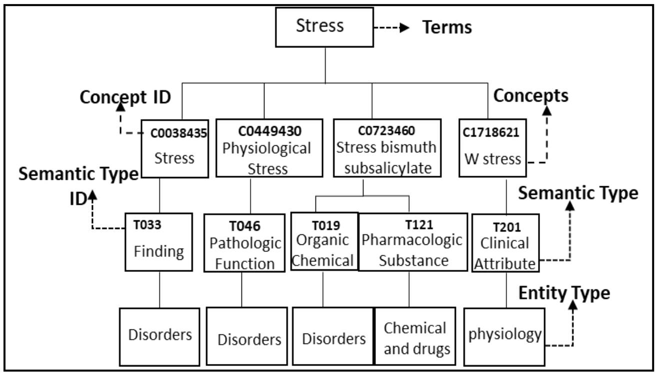

Fig. 2. Information Extraction Hierarchal Tree

UMLS terminology services are playing basic role to access UMLS Metathesaurus. We have consider a word *"Stress"* which is a medical term. We requested to Metathesaurus to identify the concept for a term *"Stress"*. Metathesaurus will return a list of concepts ID's and concept name. As shown in Fig.2 and Table I the term *"Stress"* has four concepts name extract from Metathesaurus. After concept extraction, based on concept ID's we requested to Metathesaurus to return all related semantic types for each concept name. Metathesaurus will return a list of semantic types ID and semantic types for each concept. In Fig.2 and Table I for each concept have one semantic type while the *"Stress bismuth subsalicylate"* have two semantic types *"organic chemical"* and *"pharmacological substance"*. Based on semantic type ID we requested to Metathesaurus for entity type, in response entity type ID and entity type received. The main challenge in Algo I or in the methodology is to extract a list of information, which is present in a dynamic sequence.

TABLE I
TERMS, CONCEPTS, SEMANTIC TYPES, AND ENTITY TYPES EXTRACTION

Terms	Concepts	Semantic Type	Entity Type
Stress	Stress	Finding	Disorders
	Physiological Stress	Pathologic Function	Disorders
	Stress bismuth subsalicylate	Organic Chemical	Disorders
		Pharmacologic Substance	Chemical and Drugs
	W stress	Clinical Attribute	Physiology
Blood Pressure	Blood Pressure	Pharmacological Substance	Physiology
	Blood Pressure determination	Health Care Activity	Procedures
	Blood Pressure Finding	Finding	Disorders
	Systemic arterial pressure	Finding	Disorder
Coronary Artery Disease	Coronary Arteriosclerosis	Disease or Syndrome	Disorder
	Coronary Artery Disease	Disease or Syndrome	Disorder
	Liver function tests normal	Finding	Disorder

Table I show the information extraction output of proposed algorithm in proper readable and understandable format. We have taken 4 terms with N-gram value 4 as an example.

IV. RESULTS AND DISCUSSIONS

We implement the proposed algorithm in Python, using NLTK library for pre-processing of data. In the algorithm, exact word matching procedure with UMLS Metathesaurus performed as shown in Fig.1. UMLS Metatheseaurus can be accessed by three ways provide by UMLS Terminology (a) Web Browser,(b) Local Installation and (c) Web Services API. In the current study, we are following Web Services API approach to access the UMLS Metathesaurus see Table II and Table III.

TABLE II AUTHENTICATION SERVICE END POINT

Base URL	Method Type	Path	Description
https://utslogin.nlm.nih.gov			
1	POST	/cas/v1/tickets	Retrieves a Ticket Granting Ticket (TGT)
2	POST	/cas/v1/tickets/{TGT}	Retrieves a single-use Service Ticket

TABLE III END POINTS TO SEARCH AND RETRIEVE UMLS CONTENT

Base URL	Method Type	Path	Description
https://uts-ws.nlm.nih.gov/rest			
1	GET	/search/{version}	Retrieves CUIs when searching by term or code
2	GET	/content/{version}/CUI/{CUI}	Retrieves a single-use Service Ticket
3	GET	/semantic-network/{version}/TUI/{id}	Retrieves information for a known Semantic Type identifier (TUI)

In clinical domain, most the data is present in unstructured and heterogeneous format. Our first motive is to convert the unstructured format of clinical data into structure format. Our other motive is to extract and identify the meaningful medical information from the clinical document using UMLS. We introduce a design and implementation of algorithm that extracts standard concepts, their semantic types, and entity types based on medical phrases and terms used in the clinical unstructured documents. The third motive of proposed algorithm is automatically preparation of data for driven approaches like Machine Learning because in machine learning data labeling is very expensive and time consuming which need domain expert. As a domain expert UMLS Metathesaurus is used to label, the medical concept with semantic type and entity type correctly.

We choose ten unstructured clinical notes from I2b2 2010 clinical notes test dataset for algorithm testing and evaluation purpose. In Fig.3. A bag of words generated contains words of N-gram-4 as discussed in section III in pre-processing step. We evaluated each N-gram of words, and extracted the medical information's such as terms, concepts, semantic types and entity types. After analysis, in Fig.3.we have found that UMLS Metathesaurus identify medical information for N-gram-1 more than N-gram-2 to N-gram-4. Very less information found in Metathesaurus for N-gram-4. We also realized that term having N-gram equal to or greater than 2 as discussed in section A equ.1, give more coherent and meaningful information accommodate in UMLS Metathesaurus such as “blood pressure”, “coronary artery disease” and “liver function test normal” see Table I as compared to single term such as “pressure”, “blood” etc.

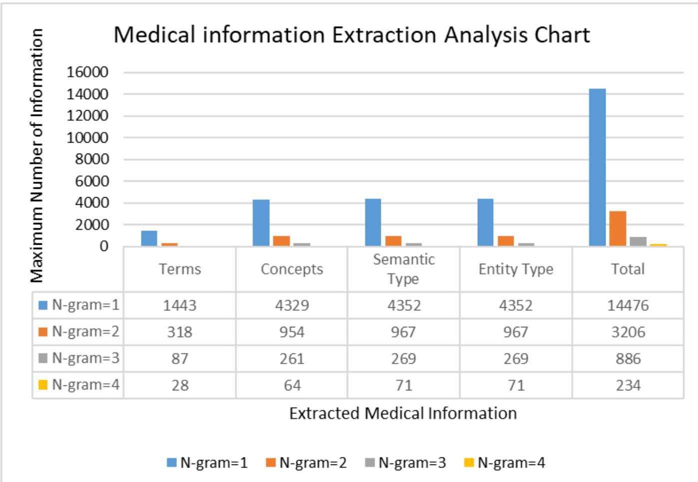

Fig. 3. Statistical Analysis Chart for Information Extraction

Table I contains four terms with N-gram size 4. As the number of N-gram size, increase the identification of composite terms in Metathesaurus decrease following exact word matching approach. Table I presenting the output of the algorithm for medical information extraction. It was a challenging task to extract a list of concept for each term, then a list of semantic type for each concept and then entity type for each semantic type. In Fig.2, we are presenting the flow of algorithm in the form tree structure because every child leaf or value is depended on the parent leaf or value. Such as first, we matched term to Metathesaurus, and then extracted a list of concept for term. Based on concept we extracted a list of semantic type for each concept. A concept have one or more than one semantic types. In the last based on semantic type, we extracted entity type for each semantic type.

V. Conclusion

In this paper, we have designed an algorithm to identify and extract a medical information from UMLS Metathesaurus followed the exact term match based approach. We evaluated the algorithm and perform medical information extraction on English clinical unstructured notes of I2b2 2010. This algorithm extracts the medical information like terms, concepts, its semantic type and entity type, which give more specific, obvious and interpretable meaning for a given term in unstructured clinical documents. The proposed algorithm could be a useful tool to annotate or label the data, which in turn could be utilized for machine learning model training and testing. In future, we will measure the efficiency of algorithm against existing tools and method available in the domain of concept extraction from patient data.

VI. Conclusion

In this paper, we have designed an algorithm to identify and extract a medical information from UMLS Metathesaurus followed the exact term match based approach. We evaluated the algorithm and perform medical information extraction on English clinical unstructured notes of I2b2 2010 clinical notes. This algorithm extracts the medical information like terms, concepts, its semantic type and entity type, which give more specific, obvious and interpretable meaning for a given term in unstructured clinical documents. The proposed algorithm could be a useful tool to annotate or label the data, which in turn could be utilized for machine learning model training and testing. In future, we will measure the efficiency of algorithm against existing tools and method available in the domain of concept extraction from patient data.

Acknowledgment

This research was supported by the Ministry of Science and ICT (MSIT), Korea, under the Information Technology Research Center (ITRC) support programme (IITP-2017-0-01629) supervised by the Institute of Information & Communications Technology Planning & Evaluation (IITP). This work was supported by an IITP grant funded by the Korean government (MSIT) (no.2017-0-00655).

References

[1] Soldaini, Luca, and Nazli Goharian. "Quickumls: a fast, unsupervised approach for medical concept extraction." MedIR workshop, sigir. 2016.
[2] Guo, Yufan, Deepika Kakrania, Tyler Baldwin, and Tanveer F. Syeda-Mahmood. "Efficient Clinical Concept Extraction in Electronic Medical Records." In AAAI, pp. 5089-5090. 2017.
[3] Divita, Guy, Qing T. Zeng, Adi V. Gundlapalli, Scott Duvall, Jonathan Nebeker, and Matthew H. Samore. "Sophia: a expedient UMLS concept extraction annotator." In AMIA Annual Symposium Proceedings, vol. 2014, p. 467. American Medical Informatics Association, 2014.
[4] Khin, Nyein Pyae Pyae, and Khin Thidar Lynn. "Medical concept extraction: A comparison of statistical and semantic methods." In 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 35-38. IEEE, 2017.
[5] Srinivasan, Suresh, Thomas C. Rindflesch, William T. Hole, Alan R. Aronson, and James G. Mork. "Finding UMLS Metathesaurus concepts in MEDLINE." In Proceedings of the AMIA Symposium, p. 727. American Medical Informatics Association, 2002.
[6] Campillos, Leonardo, Louise Deléger, Cyril Grouin, Thierry Hamon, Anne-Laure Ligozat, and Aurélie Névéol. "A French clinical corpus with comprehensive semantic annotations: development of the Medical Entity and Relation LIMSI annOtated Text corpus (MERLOT)." Language Resources and Evaluation 52, no. 2 (2018): 571-601.
[7] Tran, Le-Thuy T., Guy Divita, Marjorie E. Carter, Joshua Judd, Matthew H. Samore, and Adi V. Gundlapalli. "Exploiting the UMLS Metathesaurus for extracting and categorizing concepts representing signs and symptoms to anatomically related organ systems." Journal of biomedical informatics 58 (2015): 19-27.
[8] Aronson, Alan R., and François-Michel Lang. "An overview of MetaMap: historical perspective and recent advances." Journal of the American Medical Informatics Association 17, no. 3 (2010): 229-236.
[9] Savova, Guergana K., James J. Masanz, Philip V. Ogren, Jiaping Zheng, Sunghwan Sohn, Karin C. Kipper-Schuler, and Christopher G. Chute. "Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications." Journal of the American Medical Informatics Association 17, no. 5 (2010): 507-513.
[10] Sahu, Ruchi. "Rule-Based Method for Automatic Medical Concept Extraction from Unstructured Clinical Text." Recent Findings in Intelligent Computing Techniques. Springer, Singapore, 2018. 261-267.

[11] https://www.nlm.nih.gov/research/umls/new_users/online_learning/SEM_001.html,last accessed 2019/01/20.
[12] Segura-Bedmar, Isabel, Paloma Martínez, and María Segura-Bedmar. "Drug name recognition and classification in biomedical texts: a case study outlining approaches underpinning automated systems." Drug discovery today 13, no. 17-18 (2008): 816-823.

Asim Abbas born in Parachinar, Pakistan in 1994. He received B.S. degree in computer science from Islamia University Peshawar, Pakistan in 2015. He is currently pursuing the M.S degree with the Ubiquitous Computing Laboratory, Department of Computer Science and Engineering at Kyung Hee University, South Korea. He has professional experience for over 2 years in software industry, working on Mobile application (Android and IOS) projects. His research interest includes artificial intelligence, Text Mining, Health Informatics and Information extraction from unstructured data.

Muhammad Afzal received his PhD from Kyung Hee University, South Korea, in 2017 with Excellent Thesis Award and M.S. degree from the School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Pakistan in 2009. He is now working as Assistant Professor at Sejong University, Seoul, South Korea. He contributed to author more than 60 publications in different reputed journals/conferences in addition to 4 local and international patents. His current research includes Evidence-adaptive Clinical Decision Support Systems, Machine Learning, and Clinical Text Processing.

Jamil Hussain received the MS degree from the School of Electrical Engineering and Computer Science, National University Science and Technology, Pakistan, in 2014. He Jamil Hussain received the MS degree from the School of Electrical Engineering and Computer Science, National University of Science and Technology, Pakistan in 2014. He is currently pursuing the PhD degree with the Ubiquitous Computing laboratory, Department of Computer Science and Engineering at Kyung Hee University, South Korea. He has professional experience for over 7 years in software industry, working on user experience design and development projects. His research interest includes User experience, artificial intelligence, and Information extraction from textual data.

SUNGYOUNG LEE (M'89) received the B.S. degree from Korea University, Seoul, South Korea, and the M.S. and Ph.D. degrees in computer science from the Illinois Institute of Technology, Chicago, IL, USA, in 1987 and 1991, respectively. He was an Assistant Professor with the Department of Computer Science, Governors State University, University Park, IL, USA, from 1992 to 1993. He has been a Professor with the Department of Computer Engineering, Kyung Hee University, South Korea, since 1993, where he has been the Director of the Neo Medical ubiquitous-Life Care Information Technology Research Center since 2006. He is currently the Founding Director of the Ubiquitous Computing Laboratory. His current research interests include ubiquitous computing and applications, wireless ad hoc and sensor networks, context-aware middleware, sensor operating systems, real-time systems and embedded systems, and activity and emotion recognition. He is a member of ACM.