International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-12, October 2020

NLP Algorithms Endowed for Automatic Extraction of Information from Unstructured Free-Text Reports of Radiology Monarchy

Vaishali M.Kumbhakarna, Sonali B. Kulkarni, Apurva D.Dhawale

Abstract*: *Natural Language Processing (NLP) Algorithms are the key factors for automatic information extraction form the unstructured free-text radiology reports .To extract clinically important findings and recommendations, various NLP algorithms are used . A rule-based NLP system is used in most of the automated IE applications in medical domain; whereas some applications are using Random Forest classifier, PageRank Algorithm, clustering algorithm, Conditional Random Fields (CRF) algorithm, and deep learning-based approaches. Some papers found with methods used for IE like, Support Vector Machines (SVMs), linear-chain conditional random fields (LC-CRFs), k-means or k-medoids algorithm ,Affinity Propagation (AP) clustering algorithm , supervised machine learning algorithm and many more. Thus through this survey we can say that, NLP methods used to extract information ,brings new insights into already known clinical evidences. It also helps to identify previously unknown treatment and causal relations between biomedical entities .Therefore NLP algorithms has empowered Radiology monarchy.

Keywords : NLP,IE, EHR ,CRF,LC-CRF,AP,SVM, ,PET,MRI.

I. INTRODUCTION

Radiology is a subfield of medical science that has expanded rapidly in the year 2000 due to its non-invasive nature and advances in computer technology .Radiology discipline uses a variety of imaging techniques like X-ray , ultrasound sonography , computed tomography (CT), nuclear medicine including positron emission tomography (PET) and magnetic resonance imaging (MRI) etc.to diagnose or treatment of diseases. .The reports generated from these modalities consist of large no of unstructured free text. The modern practice of radiology involves several different healthcare professions working as a team. Circadian enormous electronic health records are generating. EHR data consist of information in both structured and unstructured data formats. This variation in data format pretence significant challenges in reform the information to be utilized. Ample of medical information is rooted in clinical histories .For instance; detailed information about patient conditions, interventions, clinical progress, and treatment outcomes are often captured in clinical notes. Natural language processing (NLP) provides openings to beat into clinical narratives to extract the information needed for various clinical applications.In this survey paper ,it is evaluated that near about 80% of the medical data exist in the unstructured layout and NLP has potential solutions, which automatically extract discrete, usable clinical data from unstructured –free text. In this study, we introduce NLP components for radiology text processing such as tokenization, syntactic parsing, semantic parsing, and pragmatic interpretation. Along with that we also include survey on speech recognition, clinical data mining, text analytics, visualization, and summarization of the NLP etc.We also take survey on a automated tool required in health care called as terminology mapping-which takes the content that is clinically relevant and produces codes for unified semantic representations of clinical concepts . These codes are subsequently used in various clinical applications such as billing, compliance, quality measurement, clinical decision support, disease-drug-treatment relationship and many more. .

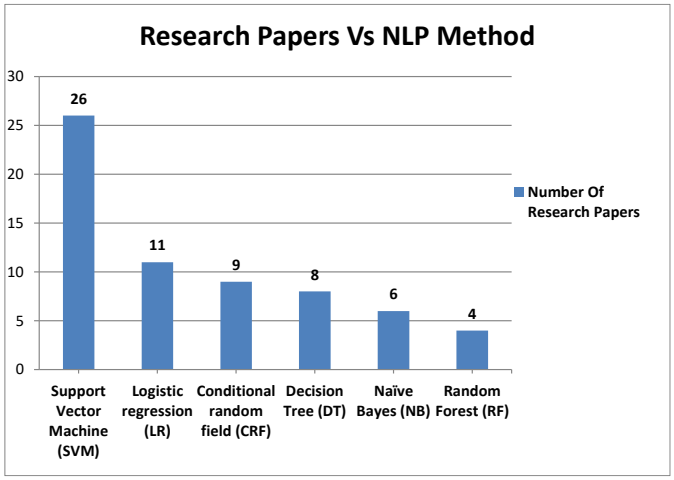
Revised Manuscript Received on October 30, 2020.
***** Correspondence Author
Vaishali M. Kumbhakarna*, Dr.Babasaheb Ambedkar Marathwada University , Department of Computer Science, Aurangabad,India. Email: vmk_17@yahoo.co.in
Sonali B. Kulkarni(Asst.Prof), Dr.Babasaheb Ambedkar Marathwada University , Department of Computer Science, Aurangabad,India. Email: sonalibkul@gmail.com
Apurva D.Dhawale, Dr.Babasaheb Ambedkar Marathwada University , Department of Computer Science, Aurangabad,India. Email: Addhawle@gmail.com

II. LITERATURE SURVEY

This Natural language processing has arisen with more and more importance in the field of radiology-sub stream of medical science. Ironic applications for health data processing are developed by using R software and python. In this paper we has taken a survey on various NLP methods used for information extraction from different types of radiology reports. In the paper maintained at refrence-1,to extract information[1] automatically 283 abdomen and 311 chest CT reports were used [28].The data is collected from 50 patients. In pre-processing ,segmentation is applied and from radiology reports only measurements part are extracted. Then by using binary classification ,total thirteen features were extracted. After feature extraction step, Random Forest classifier was applied to integrate all features.In post processing technique ,all 13 features are categorized into 4 groups. The mutual best match filtering mechanism is used to find out partial uniqueness within two consecutive reports and thus construction-pairing of reports-extracted information was generated automatically by using machine-learning algorithum. In the paper maintained at refrence-2 author has analysed the material of publications within the medical Natural Language Processing domain. Author found various methods and application of NLP [2] in medical text processing are given bellow.

Sr.No.	Method	Medical Monarchy
1	A rule-based NLP system	Domain-specific ontology information retrieval from medical reports of patients with arrhythmias
2	Unified Medical Language System	A set of files and software that brings together many health and biomedical vocabularies and standards to enable interoperability between computer systems
3	MetaMap	A tool for identification of UMLs concept in text and used as non-invasive biomarkers for recognition of "pre-clinical" Alzheimer disease
4	PubMed using PageRank Algorithm	A tool to detect novel and emerging drug terms.
5	Topic- based sentence clustering algorithm	A summarizer tool to extracts biomedical concepts from the input documents and employs an Item set mining algorithm to discover main topics
6	CIBS method	A tool for single and multi-document biomedical text summarization.
7	rule-based Conditional Random Fields (CRF) algorithm, and deep learning	A tool to extract BIRADS finding for breast cancer radiology reports.
8	Support Vector Machines (SVMs) and deep learning .	A tool for drug safety surveillance in electronic health records.

In the paper "An enhanced CRFs-based system for information extraction from radiology reports", [3],author studied set of 500 free-text mammography reports [30]using supervised machine learning algorithm for information extraction(IE) based on linear-chain conditional random fields (LC-CRFs) method. In the paper "Automatic Extraction of Major Osteoporotic Fractures from Radiology Reports using Natural Language Processing" , [4], author had documented a rule-based NLP algorithm for automatic extraction of six major osteoporotic fractures from radiology reports. In the paper "A bibliometric analysis of natural language processing in medical research", [5] , author conducted pre-processing using software R for dataset analysis ,and had developed Python program which cover Affinity Propagation (AP) clustering algorithm and the dataset was analysed for statistical characteristics.In addition , author defined keywords and PubMed medical subject headings (MeSH), key terms extracted from title and abstract fields included in AP and also compared with k-means and k-medoids algorithm. In the paper, "Toward a Learning Health-care System – Knowledge Delivery at the Point of Care Empowered by Big Data and NLP" [6],author studied the clinical NLP systems - the Medical Language Extraction and Encoding System (MedLEE). It was developed on chest radiology reports. In the paper," Clinical information extraction applications: A literature review" , [7] review of recent published research on clinical information extraction (IE) applications was studied. Since 2009 to 2016 total of 1917 publications were identified for title and abstract screening. Out of which, 263 articles were selected and discussed for clinical IE tools, methods, and applications in the areas of disease and drug related studies, and clinical workflow optimizations .In this review,author found 26 research papers using Support Vector Machine (SVM) algorithm, 11 papers used Logistic regression (LR) method,Conditional random field (CRF) had studied in 9 papers where as few researchers were used Decision Tree (DT), Naïve Bayes (NB) and Random Forest (RF) algorithm for studies.



In the paper," Clinical Natural Language Processing with Deep Learning", [8].Author quoted that ,to classify radiology free text reports based on pulmonary embolism findings, CNN-based models have shown better performance over the traditional machine learning classifiers .For automated coding of radiology reports[36] using the International Classification of Diseases (ICD-10) coding scheme, RNN architecture was better and for clinical paraphrase generation CNNs are generally shown to be effective in solving classification tasks such as sentiment analysis, spam detection, or topic categorization .CNN is a multilayer neural network that uses a special kind of linear mathematical operation. RNNs generally work well for modeling sequences. Hence, they are used to solve various NLP tasks due to their ability to deal with variable-length input and output. In the paper," Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification" , [9] author proposed a innovative idiom attention-based hierarchical recurrent neural network model (DPA-HNN) that can accurately classify free text chest CT reports into pre-defined PE related criteria. NLP methods suggest feasibility of CNNs and RNNs in automated classification of imaging text reports and support the application of these techniques at scale in classifying free text imaging reports for various use cases including radiology patient prioritization, cohort generation for clinical research, eligibility screening for clinical trials, and assessing imaging utilization had studied. In the paper," Design and Development of a Multimodal Biomedical Information Retrieval System", [10] author had discussed about the Multimodal biomedical information retrieval system named as "OpenI".it is used in biomedical domain for retrieval of case descriptions similar to a patient's case in the form of images and text data[29]. The OpenI document processing system is developed in Java, and uses Hadoop. The rule-based algorithm is used for text processing where as features extracted from the images were clustered using the k-means algorithm, The OpenI prototype supports image retrieval for textual, visual and hybrid queries. a Naïve Bayes classifier, and classifiers based on the position of the sentence in the abstract, or the presence of relevant terms in the sentence was used.

Thus increasing commercial interest of researchers in multi-modal information retrieval in the biomedical domain is used for retrieval of medical images and text about similar patients' cases .

In the paper," Application of Recently Developed Computer Algorithm for Automatic Classification of Unstructured Radiology Reports: Validation Study" [11] author demonstrated the potential benefits of structured medical reports for research, teaching, and organization of patient medical records from the electronic radiology report database. Author studied LEXIMER search engine with data of 1119 radiology reports obtained from diffrent imaging modality listed in the table given bellow.

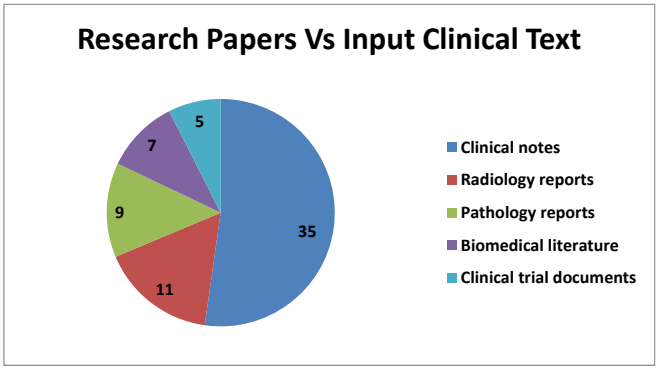
Sr.No.	Imaging Modality
1)	Barium Studies
2)	Computed Tomography (CT)
3)	Mammography
4)	Magnetic Resonance Imaging (MRI)
5)	Nuclear Medicine
6)	Positron Emission Tomography (PET)
7)	Radiography
8)	Ultrasonography (US)
9)	Vascular Procedures

LEXIMER categorize the reports into FT and FT0 (containing or not containing clinically important findings) and RT and RT0 (containing or not containing recommendations for subsequent action) by using stemming algorithm. In the paper maintained at refrence-12 author had developed the system named as,BANNER [12].It was a trainable NER system built on linear-chain conditional random fields. The frame contains 4 kinds of clinical notes: discharge summaries, ECG, 2D echo , and radiology reports. This system has prospective such as clinical decision support[38] , used for patients appreciative records , for public health care like. Bio-surveillance, and in biomedical research . In the paper," Assigning clinical codes with data-driven concept representation on Dutch clinical free text", [13] ,author had developed algorithms for automated code generation of clinical procedure.Uunsupervised and semi-supervised methods for the extraction of multi-word expressions that convey a generalizable medical meaning were used.For this study two types of datasets were derived .The first dataset consists of an un annotated corpus .The second dataset consists of a randomized subset of annotated patient data with associated documents (radiology reports, requests, surgery reports, notes, letters, and attestations) The dataset was divided into three specialties (cardiology, gastroenterology, and urology) .The datasets were pre -processed with several low-level natural language processing steps (sentence splitting, tokenization, lemmatization, part-of-speech tagging, and chunk tagging) .

In the paper, "Crowd control: Effectively utilizing unscreened crowd workers for biomedical data annotation ", [14], author examine the effectiveness of crowd sourcing with unscreened online workers as an alternative for transforming unstructured texts in EHRs.[35] The annotated data were directly usable in supervised learning models .Author had conducted a retrospective analysis of de-identified radiologist reports .

In the paper,"Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review "[15].In this review ,author studied many papers worked on Rule-based NLP method,and few were used hybrid and machine learning NLP.The most common form of input text processed is listed bellow.

Sr.No.	Input Clinical Text
1)	Clinical notes or prescriptions
2)	Radiology reports
3)	Pathology reports
4)	Biomedical literature
5)	Clinical trial documents



In this study,author observed ,that various programming language were used for text processing,like Java ,perl,python and R software and maximum research was done on text processing by using clinical notes. In the paper,"Selecting relevant features from the electronic health record for clinical code prediction", [16]author had developed a Computer-assisted coding prediction system for clinical codes genetation.Clinical codes reflect diagnoses and procedures related to a patient stay and are primarily assigned for reporting and reimbursement purposes. Techniques used for feature selection, was markov-blanket algorithms. In the paper," Deep neural models for ICD-10 coding of death certificates and autopsy reports in free-text", [17]author had presented the deep neural network model for assigning ICD-10 clinical codes to underlying causes of death, by analysis of the free-text information within death certificates. The associated post-mortem or autopsy reports and clinical bulletins, from the Portuguese Ministry of Health were studied. Automatic assignment of ICD codes to clinical text was a computational medicine challenge. This system was used for real-time surveillance of specific causes of death. Recurrent Neural Networks text classification was used for it.

In the paper," Recognition and Evaluation of Clinical Section Headings in Clinical Documents Using Token-Based Formulation with Conditional Random Fields" ,[18] author had developed.

Token-Based formula by using Conditional Random Field algorithum to recognize section heading from clinical documents. For standard database of discharge summaries author observed that each country adopted different clinical section heading. So in this study he used the standard format of clinical documents defined by the electronic medical record exchange centre built by the Ministry of Health and Welfare of Taiwan. This research is useful for Individual patient or population care, clinical research, quality improvement, and public health. For instance, the use of EHRs [37]eliminates the manual task of extracting data from charts and also promotes the access, retrieval, and sharing of clinical information.

In the paper," Biomedical/clinical NLP", [19] , author had given overview of the biomedical and clinical NLP data, tools, and methods for building annotated corpora and also discussed about linguistic aspects of data; syntactic: part of speech tagging, parsing and semantic :concept extraction, temporal information extraction, co reference resolution.The NLP methods reviewed in this study were rule-based, statistical, and hybrid methods.

In the paper, "General Natural Language text Processing for Clinical Rediology" [20].author had developed a semantic based text processor that extracts and structures clinical information from textual radiology reports and translates the information to terms in a controlled vocabulary so that the clinical information can be accessed by further automated procedures.

In the paper ,"Natural language processing to identify ureteric stones in radiology reports", [21]author had discussed the concept of automated data extraction with NLP . Data was summarized in two main steps.Step one: Identification of concepts achieved by pattern matching and linguistic analyses. Step two: Determination of whether the output data contain the desired concepts and modifiers required for a particular purpose. This was achieved using a set of clinical logic rules or statistical and machine learning methods.Author also cited that , research into NLP utilization in radiology domain can be divided into five main categories.

Sr.No.	Radiology Category	Purpose
1)	Diagnostic surveillance	Automated detection of critical findings requiring follow up
2)	Unit building for epidemiological studies	Slection of appropriate cases positive for a certain condition
3)	Query-based case retrieval	Identification of cases with conditions that are not predefined but specified in a query
4)	Quality assessment of radiological practice	Generating descriptive statistics on recommendation , completeness and communication of critical results
5)	Clinical support services	Integrated applications into clinical workflow of radiologists

For this study word normalization algorithms used including stemming, spellcheck and expansion of abbreviations.Author also concluded that syntax recognition was another common problem due to the variation of radiologist in describing single medical term in many ways. In the paper, maintained at refrence-22,author studied breast cancer cases[22]. Identified probable cases can be easily fixed by studying the noted reports .The TIES system was a great tool for identifying prospective cancer cases [31]in a well-timed mode .The studies had been focused on developing methods for semantic processing and analysis of clinical texts.

In the paper, "Making Sense of Big Textual Data for Health Care: Findings from the Section on Clinical Natural Language Processing", [23]author had reviewed best papers of the year 2016 in the field of clinical NLP. Author concluded that Machine Learning Algorithm and supervised classifiers were used in many clinical text processing.

In the paper, "TEXT2TABLE:Medical Text Summarization System based on Named Entity

Recognition and Modality Identification", [24]author had proposed a system that extracts medical events from unstructured text of discharge summary and convert it into a table structure. Conditional Random Fields (CRFs) method was used for information extraction and SVM-based classifier was used for syntactic information extraction to find Negative events.

In the paper, "Natural language processing: State of the art and prospects for significant progress", [25] author cited that NLP is vital for innovative healthcare .It is essential to convert applicable data locked in text into well-structured format and thus can be utilize for improving patient care [27]. Author cited two major NLP policies :1) symbolic method

2) statistical method. the most effective systems may combine the two approaches.

In the paper,"Machine Learning and Radiology ",[26] author gave survey on six diffrent categories of machine learning applications in radiology listed in the table below.

Sr.No.	Machine Learning Applications Zone In Radiology
1)	Medical Image Registration
2)	Medical Image Segmentation
3)	Computer Aided Detection And Diagnosis
4)	Brain Function Or Activity Analysis And Neurological Disease Diagnosis From FMRI
5)	Content-Based Image Retrieval Systems For CT Or MRI
6)	Text Analysis Of Radiology Reports Using Natural Language Processing (NLP) And Natural Language Understanding (NLU)

In this study author came across various methods like, supervised learning, unsupervised learning ,Semi-supervised learning, Artificial neural networks (ANNs) , Cellular neural network (CNN) ,Support vector machines (SVMs) ,Bayesian network [33],Markov network ,Bagging algorithm ,k-means clustering ,Fuzzy c-means clustering, Principal Components Analysis (PCA), independent component analysis (ICA) ,k-nearest neighbor (k-NN) classifier etc. In the paper," Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications", author developed the Medical Language Extraction and Encoding System (MedLEE) . It was the first clinical NLP systems developed for chest radiology reports. IBM Watson, IBM has invested in health-care analytics clinical Text Analysis and Knowledge Extraction System (cTAKES) [32] built on the Unstructured Information Management Architecture (UIMA), Med Tagger was tool developed for indexing medical concepts, information extraction, and named entity recognition,

MedXN developed for medication extraction and concept normalization, MedTime for clinical temporal information extraction, are some of the NLP tools developed for clinical realm.

III. CONCLUSION

In this survey paper, we had proven that, range of NLP methods are associated with building a successful automatic information extraction system based on free-text radiology reports .This survey concluds that by using NLP algorithms many clinical decision support systems are developed and used in real time.SVM classifier is used in most of the two case classification like diagnosis of positive or negative patient investigation .The clinical decision support based on findings can automatically diagnose the patient report.CNN is used in AI and machine learning for automatic detection of any disease.Thus in this survey,we had verified that artificial intelligence used through NLP has empowered the medical science and is important factor of modern radiology realm .NLP is efficient for clinical text processing and IE from clinical notes,priscription,test reports ,discharge card summery etc. and also less time consuming ;yes but for authorization and authentication of automatic diagnoses we can't deny the need of expert radiologist.

REFERENCES

1. Merlijn Sevenster , Jeffrey Boze,man, USA," A natural language processing pipeline for pairing measurements uniquely across free-text CT reports" ,Journal of Biomedical Informatics , 1532-0464 ,Elsevier ,2015.
2. Natalia Grabar, Cyril Grouin, France," A Year of Papers Using Biomedical Texts:Findings from the Section on Natural Language Processing of the IMIA Yearbook", IMIA Yearbook of Medical Informatics ,2019.
3. Andrea Esuli, Diego Marcheggiani, Italy," An enhanced CRFs-based system for information extraction from radiology reports" , Journal of Biomedical Informatics,Elsevier ,2013.
4. Yanshan Wang, Saeed Mehrabi ,USA," Automatic Extraction of Major Osteoporotic Fractures from Radiology Reports using Natural Language Processing" , International Conference on Healthcare Informatics Workshop,978-1-5386-6777-4 ,IEEE ,2018.
5. Xieling Chen, Haoran Xie, China," A bibliometric analysis of natural language processing in medical research", Medical Informatics and Decision Making ,s12911-018-0594-x, BMC,2018.
6. Vinod C. Kaggal, Ravikumar Komandur Elayavilli, USA." Toward a Learning Health-care System – Knowledge Delivery at the Point of Care Empowered by Big Data and NLP" Biomedical Informatics Insights, 2016.
7. Yanshan Wang, Liwei Wang, United States," Clinical information extraction applications: A literature review", Journal of Biomedical Informatics , Elsevier, 2018.
8. Sadid A. Hasan , Oladimeji Farri," Clinical Natural Language Processing with Deep Learning" ,Data Science for Healthcare, Springer, 2019 .
9. Imon Banerjeea, Yuan Lingb, USA," Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification", Artificial Intelligence In Medicine,0933-3657, Elsevier, 2018 .
10. Dina Demner-Fushman, Sameer Antani, USA," Design and Development of a Multimodal Biomedical Information Retrieval System" Journal of Computing Science and Engineering, Vol. 6, eISSN: 2093-8020, 2012.
11. Keith J. Dreyer ,Mannudeep K. Kalra," Application of Recently Developed Computer Algorithm for Automatic Classification of Unstructured Radiology Reports: Validation Study" , RSNA ,Radiology; 234:323–329, 2005.
12. Robert Leaman, Ritu Khare, United States ,"Challenges in clinical natural language processing for automated disorder normalization" ,Journal of Biomedical Informatics ,2015.
13. Elyne Scheurwegs , Kim Luyckx, Belgium," Assigning clinical codes with data-driven concept representation on Dutch clinical free text", Journal of Biomedical Informatics, Elsevier ,2017.
14. Anne Cocos , Ting Qian ,United States, "Crowd control: Effectively utilizing unscreened crowd workers for biomedical data annotation ", Journal of Biomedical Informatics, Elsevier ,2017.
15. Kory Kreimeyer , Matthew Foster , United States ,"Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review", Journal of Biomedical Informatics , Elsevier ,2017.
16. Elyne Scheurwegs , Boris Cule ,"Selecting relevant features from the electronic health record for clinical code prediction", Journal of Biomedical Informatics 74, 92–103, Elsevier,2017.
17. Francisco Duartea,∗, Bruno Martinsa," Deep neural models for ICD-10 coding of death certificates and autopsy reports in free-text", Journal of Biomedical Informatics , Elsevier, 2018.
18. Hong-Jie Dai,1,2 Shabbir Syed-Abdul," Recognition and Evaluation of Clinical Section Headings in Clinical Documents Using Token-Based Formulation with Conditional Random Fields", BioMed Research International, 2015.
19. Ozlem Uzuner," Biomedical/clinical NLP", International Conference on Computational Linguistics, 2014.
20. Carol Friedman PHD,Philip Alderson MD,"General Natural Language text Processing for Clinical Rediology", Journal of the American Medical Informatics Association , 1994.
21. Andrew Yu Li1, Nikki Elliot, New Zealand ,"Natural language processing to identify ureteric stones in radiology reports", Journal of Medical Imaging and Radiation Oncology ,2019.
22. Fagen Xie, Janet Lee, USA," Application of Text Information Extraction System for Real-Time Cancer Case Identification in an Integrated Healthcare Organization", Journal of Pathology Informatics ,2017.
23. A. Neveol, P. Zweigenbaum, "Making Sense of Big Textual Data for Health Care: Findings from the Section on Clinical Natural Language Processing", IMIA Yearbook of Medical Informatics, 2017.
24. Eiji ARAMAKI ,Yasuhide MIURA ,"TEXT2TABLE:Medical Text Summarization System based on Named Entity Recognition and Modality Identification",Proceedings of the Workshop on BioNLP,Association for Computational Linguistics, 2009.
25. Carol Friedman , Thomas C. Rindflesch, "Natural language processing: State of the art and prospects for significant progress, a workshop sponsored by the National Library of Medicine" , Journal of Biomedical Informatics 46, 765–773 ,Elsevier, 2013.
26. Shijun Wang and Ronald M. Summers,"Machine Learning and Radiology ",Med Image Anal, 2012.
27. Ewoud Pons," Natural Language Processing in Radiology: A Systematic Review", RSNA, 2016
28. Imon Banerjee," Radiology report annotation using intelligent word embeddings: Applied to multi-institutional chest CT cohort", Elsevier, 2017.
29. Pragya A. Dang," Informatics in Radiology Render: An Online Searchable Radiology Study Repository", RSNA, 2009.
30. Tianrun Cai," Natural Language Processing Technologies in Radiology Research and Clinical Applications" , RSNA, 2016.
31. Iain Mccowan," Automated Cancer Stage Classification from Free-text Histology Reports", ResearchGate, 2006.
32. Guergana K Savova," Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications", jamia,2009.
33. Fuchiang Tsui," Probabilistic Case Detection for Disease Surveillance Using Data in Electronic Medical Records", Online Journal of Public Health Informatics ISSN 1947-2579 ,2011.
34. Naveed Afzal," Mining peripheral arterial disease cases from narrative clinical notes using natural language processing", Journal of Vascular Surgery , ResearchGate,2017.
35. V. G. Vinod Vydiswaran," Special issue of BMC medical informatics and decision making on health natural language processing", BMC Medical Informatics and Decision Making, 2019.
36. Imre Solti," Automated Classification of Radiology Reports for Acute Lung Injury: Comparison of Keyword and Machine Learning Based Natural Language Processing Approaches",IEEE,2009.

37. Simon Kocbek," Text mining electronic hospital records to automatically classify admissions against disease: Measuring the impact of linking data sources", Elsevier ,2016 .
38. Dina Demner-Fushman," What can natural language processing do for clinical decision support? ", Elsevier ,2009.

AUTHORS PROFILE

Ms.Vaishali M.Kumbhakarna Master of Science from Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, India with First class in year 2013. And M.C.M. from Pune University in 2001.She has also completed M.Phil in Computer Science from Dr.BAMUniveristy, Aurangabad in the year 2016 and Currently she is pursuing her Ph.D. in Computer Science from Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, India.currently working as Assistant Professor in Department of Computer Science and IT, M.G.M University, Aurangabad. since 2008. She has 12 years of teaching experience. She has published 8 research papers in reputed international journals including Scopus, Elsevier, Springer. Her main research work focuses on, biomedical digital image processing and Natural Language Processing

Dr.Sonali B.Kulkarni Master of Science from Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, India with First in the order of merit in year 2002.She has also completed Ph.D in Computer Science from Dr.BAMUniveristy, Aurangabad and currently working as Assistant Professor in Department of Computer Science and IT, Dr.BAM University, Aurangabad since 2008. She has 18 years of teaching and Research Experience.She is a life member of IETE, ISCA, Grant and IJENG She has published more than 30 research papers in reputed international journals including Scopus, Thomson Reuters, ICI, Web of Science and conferences including IEEE, Elsevier, Springer. Her main research work focuses on Remote Sensing & GIS, Brain Computer Interface, Natural Language Processing and Linguistics.

Ms. Apurva D. Dhawale completed M.phil in Computer Science in 2015 from Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, India. Currently she is pursuing her Ph.D. in Computer Science from Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, India. She has 9 years of teaching experience, and published 9 papers in international & national reputed Journals/Conferences till date. Her research interest areas are Natural Language Processing & Biometric Image Processing.