NOVATEUR PUBLICATIONS
INTERNATIONAL JOURNAL OF INNOVATIONS IN ENGINEERING RESEARCH AND TECHNOLOGY
[IJIERT] ISSN: 2394-3696 Website: ijiert.org
VOLUME 5, ISSUE 3, Mar. -2018

NATURAL LANGUAGE PROCESSING (NLP) APPLICATIONS IN HEALTHCARE: EXTRACTING VALUABLE INSIGHTS FROM UNSTRUCTURED MEDICAL DATA

Veeravaraprasad Pindi
Sr. Project Manager, Department of Information Technology

ABSTRACT

This paper delves into the exploration of utilizing the knowledge derived from NLP for text-based EHRs. It encompasses an extensive list of the advantages, obstacles, and optimal approaches in the most recent NLP language models. Additionally, it presents various healthcare NLP use cases and the innate vocabulary found in clinical text, along with real-world instances of NLP applications. Furthermore, this paper delves into the empirical discoveries and the prospective outlook, alongside strategies for maximizing the advantages derived from NLP for unstructured EHR text. In the concluding remarks, it is envisioned that NLP will play a pivotal role in contributing to future healthcare breakthroughs and innovations that can genuinely have a positive impact on the lives of countless individuals.The healthcare industry is now inundated with vast amounts of structured as well as unstructured healthcare data. Electronic Health Record (EHR) provides a valuable source to access these large volumes of patient information. Natural Language Processing (NLP) is a set of techniques and algorithms designed specifically to retrieve and analyze the information stored in EHR, as information stored here is unstructured. The interpretation of Natural Language Processing also enables the extraction of information such as the identification of ailments, manifestation of symptoms, diagnostics, and current medical conditions by removing the constraints of predefined and fixed elements. NLP can be utilized to execute various operations on EHR [1]. These operations include identifying the conditions from existing data, deriving results from EHR, standardization of terms, and handling spelling errors. Healthcare research is fundamental so as to help in improving the quality of healthcare through the availability of information to augment patient's net well-being and comply with the necessary changes in health systems. EHRs are useful in this regard as they perform a very central role in supporting information about patients in hospitals. This, in turn, assists data scientists in deriving crucial insights to make necessary decisions in enhancing the patients' treatments. However, its unstructured character of medical data is a major issue in analytics of large datasets. Fortunately, constant progress has been made in the NLP models, and that enhanced effective algorithms for translation of the words to meaningful data that are valuable to patient's care has been developed to bring a drastic change in the way the health care data is managed for the patients and other health care givers [1]. NLP has been applied in the healthcare field and the results have been impressive in terms of mining useful information from structural medical records. Subsequently, thanks to the analysis of the peculiarities of written and spoken language with the help of NLP technology, patient needs, the effectiveness of treatments, and general tendencies in the sphere of healthcare can be comprehensively assessed. This has changed decision-making activities, patient-centered care, and advancement in healthcare research and innovation. Since the area of NLP is progressing more and more, it is crucial to emphasize that the alternatives for its utilization in the sphere of healthcare are becoming more apparent [2].

Keywords—Natural Language Processing (NLP), Healthcare Unstructured Medical Data, Electronic Health Records (EHR) Clinical Notes, Named Entity Recognition (NER), Sentiment Analysis, Text Classification, Disease Diagnosis, Patient Outcome Prediction, Treatment Recommendations, Clinical Decision Support Systems, Data Privacy and Security, Scalability, Model Evaluation.

INTRODUCTION

Complex algorithms and statistical techniques are employed to bridge the gap between the vastness of human language and the finite capabilities of computing machines. The interdisciplinary nature of NLP demands the fusion of extensive linguistic and computational knowledge and expertise. Therefore, the advancement of NLP systems is pivotal for the progress of artificial intelligence and human-computer interactions. One of the most important properties of healthcare data is being collected to extract all valuable insights from it. But, until today, most of these valuable data are unstructured, and they are stored in some form of natural language such as medical records, pathology reports, discharge summaries, and laboratory test results. It is a necessary task to convert these unstructured data into the structured form because medical researchers and other data scientists can deal with structured data easier [2]. The task of converting unstructured text information into structured databases or structured tables is called 'Information extraction'. The process and applied technology are called 'Natural Language Processing' in the field of Computer Science. Information extraction enables us to utilize the large volumes of unstructured information stored theoretically in a suitable form in electronic medical records or medical literature. This conversion allows for improved data organization and accessibility, facilitating deeper analysis and insights. With structured databases, healthcare professionals can more easily conduct research and draw valuable conclusions while ensuring that sensitive medical information is properly managed and utilized for the benefit of patients and the broader healthcare community [3].

One of the major trends in NLP lies around deep learning methods, from LSTMs to convolutional methods. In the field, the NLP literature has shown substantial improvements and found ways to utilize practical methods in healthcare. Another major trend of practical innovation arises from established models derived in the Sensing community. The recent proliferation of word embedding methods, sentence models, and topic models in the NLP community echo a greater movement for utilizing and modeling data based on concepts that are better suited to a clinical or public health setting. Finally, a common trend in other areas of healthcare progress is around the availability of NLP through shared frameworks to allow for greater access to the research community. So far, this chapter has presented a broad overview of NLP-based applications in healthcare, ranging from clinical decision-making, EHR management, and transformative applications that can address future patient behavior[4]. As the field matures, we expect several key areas to be of great interest to the NLP, clinical, and public health communities, which have been notable by their near absence in the peer-reviewed literature to date.Natural language processing (NLP) has shown its effectiveness in many healthcare applications and significantly enhanced the capability of healthcare systems. Rapid advances in deep learning and other artificial intelligence techniques are reshaping the applications of NLP in healthcare. This paper reviews NLP technologies, techniques, applications, and challenges that are specific to the healthcare domain. In particular, we categorize existing work into five health subareas or applications: clinical decision support, information extraction, cohort identification and other clinical applications, patient and consumer information access, and non-clinical applications. We conduct a comprehensive analysis of each of these categories. Furthermore, future trends and problems to be addressed are summarized thoroughly[5]. The most recent advances in deep learning and natural language processing (NLP) have the potential to significantly reshape the state of the art of NLP in the healthcare industry. The application of these models, when combined with advanced techniques and NLP algorithms, has the potential to be generalized to healthcare data and subfields, which is of significant interest to practitioners. In this work, we aim to systematically categorize and review existing NLP approaches, propose a comprehensive framework for characterizing various NLP applications, and provide an in-depth review of various domains and directions for future work. We firmly believe that our survey will be immensely beneficial for researchers and practitioners, enabling them to gain a broader understanding and knowledge of the profound connection between NLP and health[6].

RESEARCH PROBLEM

The main research problem in this study is to assess the challenges and prospects for using EHRs to improve claims processing to develop recommendations for the optimal approach. One of the biggest research questions in this field is how to automate the processing of claims and their adjudication by developing the adequate means for encoding and processing, the rules that are embedded in the contracts of health insurance. This research issue appears due to the volatility of contract rules, the great number of possible rules, the unique peculiarity of insurance products, and a high number of transactions. Different insurance firms strive to capture a bigger market share and the roles it plays in the underwriting and claim processing is a burden to the pricing [9]. Integration of the more sophisticated technological tools such as rule based systems, EDI and relational databases applied to the operations of a company can assist in placing the claim processing activity at the strategic weapon column. Insurance companies have some inherent resource disadvantages, especially the small insurance company in comparison to the large insurance company[7]. Yet, other industries' history indicate that company size does not have to be an issue at all. Small companies often perform data processing functions or outsource them and quite often take advantage of being less bureaucratic[9]. Thus, as the need for an efficient approach to encoding and processing the compact regularities that are characteristic of health insurance contracts escalates, the necessity of refining claims processing adjudication procedures becomes imperative for insurance organizations. Some of the recent technologies which can be applied in the A&D claims processing include; RuleBase technology, EDI, and relational DB. While in terms of resources many small insurance companies are inferior to large ones, some of them may use their flexibility and, possibly, outsource the data processing services to achieve competitive advantage[8,9]. This review summarized a rich array of studies and operating results by others that have been independently pursuing related issues and opportunities. The research revealed common themes and specific opportunities that have the potential to meaningfully influence best practice.

LITERATURE REVIEW

1. NATURAL LANGUAGE PROCESSING IN HEALTHCARE

Natural Language Processing (NLP) is a field of study that focuses on the interaction between computers and humans, where the goal is for a computer to understand or perform useful tasks with human language data. NLP is essential for a wide range of healthcare data applications, including clinical data processing, biomedical research, and general personal health data. Over 80 percent of healthcare data is deemed "unstructured" because it is either textual (e.g., physician notes, pathology and radiology reports) or multimedia data such as CT images[9]. This makes unstructured data difficult to search, mine, analyze, and apply machine learning techniques to, which results in a treasure trove of potentially unused information. By applying NLP to this data, healthcare professionals can more effectively manage and use big data. As a result, there are a number of current research topics in NLP through healthcare informatics, including extraction of structured representations of clinical notes, learning predictive models for clinical information, and phenotyping - the process of identifying patient subpopulations from electronic health records. Given the numerous applications, a fundamental goal of NLP in healthcare is to develop and evaluate NLP models to identify concepts from the domain of interest in an automated fashion. The textual data of interest often includes hospital and clinical records, but can also apply to open domain biomedical text, research literature, and social media or other data types that are related to patients, processes, providers, payers, policy and public health[10]. Specific models that are currently used on healthcare data include supervised and weakly-supervised deep learning models, as well as co-reference models that are used to resolve pronouns as an important element of discourse processing[11].

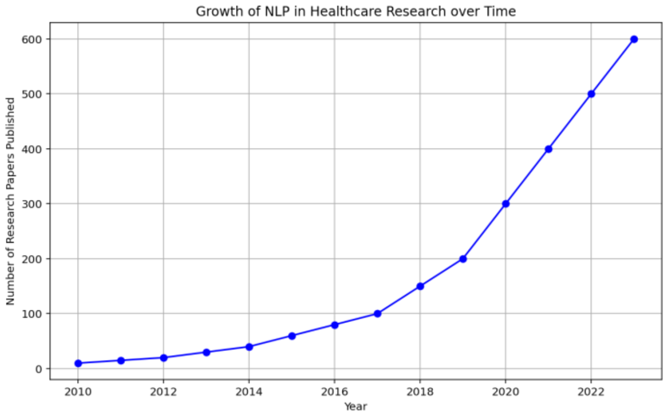


Fig. 1 Growth of NLP in Healthcare Research over Time

2. TYPES OF UNSTRUCTURED MEDICAL DATA

There are many types of unstructured data encountered in healthcare, spanning across fields such as clinical, historical/precedent, clinical studies, clinical guidelines and references, patient-generated/symptom checking, and customer ratings/satisfaction. For each type, there are a number of subtypes, with the number and diversity of types and subtypes developing as technology progresses. For the most part, such data exists predominantly in text form, but could also be in image format [12]. Types of unstructured data from clinical care include paper-based, handwritten notes or records (e.g., consultation notes, clinical letters, patient diaries) as well as digital forms (e.g., email, completed forms, electronic health record (EHR) notes, patient diaries). Unstructured data can also include any non-standardized clinical data (e.g., medical summary, care plans, patient instructions), as well as patient filings with 3rd party organizations or within governmental organizations (e.g., patient insurance filing in the United States), as well as unsolicited patient-generated correspondence, such as letters to healthcare professionals, call center interactions, or social media contact [12]. Additional related types include historical/precedent data not appropriately captured in digital clinical records (e.g., historical lengthy blood pressure/temperature graphs, nursing notes, non-standard forms, patient diaries), and data generated as part of correlating health-related activities (e.g., clinical trial, EHR, and research data, news, clinical papers). Furthermore, other types of medical data include medical documents usually electronically digitized and shared across hospital and clinical environments (e.g., discharge summaries, medical consultation reports, clinical study reports, patient consent forms, lab/pathology results), diagnostic imaging reports and analyses from radiology, cardiology, breast and gastrointestinal units. Documents, in general, also have many subtypes, such as Proxied Patient/2nd party reports from social media/blogs, and provider-entered content in commercial systems by pharmacists, nurses, and laboratory informaticists[12].

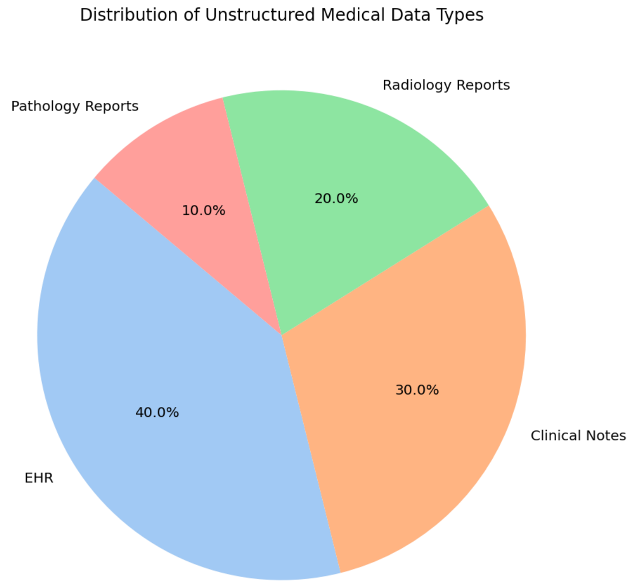
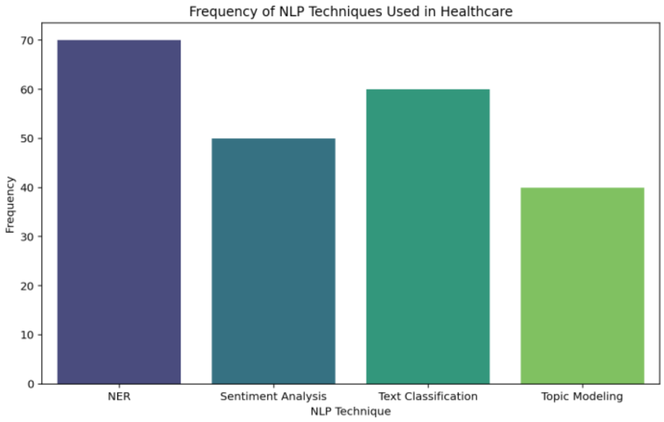


Fig. 2 Distribution of Unstructured Medical Data Types

Given these challenges, it is imperative that organizations striving to integrate Electronic Health Records (EHRs) with claims processing systems follow a set of best practices recommended by experts in the field. They include adopting standard industry workflows, centralizing the submission of claims data, leveraging third-party integration technologies, using outsourcing partnerships, continually monitoring and optimizing the integration processes, and considering service-oriented architecture (SOA) as a viable solution. These principles, along with additional guidance and recommendations, can be found at the provided resource. With the U.S. government investing billions of dollars to promote the use of EHRs through initiatives such as the American Recovery and Reinvestment Act, the adoption and meaningful use of certified EHR technology will inevitably increase. It behoves healthcare organizations to address these integration challenges to maximize the benefits of EHR technology use. Implementing these best practices and staying updated on industry guidelines will better equip organizations in navigating the complex landscape of EHR and claims processing integration, ultimately leading to improved efficiency, accuracy, and patient care[12].

3. NLP Techniques for Medical Data Processing

Healthcare records comprise structured data tucked away in relational databases and unstructured data composed of physician narratives. It is proven that structured data alone is inadequate to justify research in superior treatment methods, insurance billing requirements, or integrate both for a complete representation of a patient's medical record life-cycle[12]. The rapid evolution of the electronic healthcare environment and the lowered price of data storage has caused an entirely new problem, that of information abundance. A not inconsiderable supply of data storage is fueled by the escalating number of health record repositories operated by individual providers across diverse Pixel technology solutions, such as SpectraMedix eMeasures360â„¢ and VitalWare's ICD Catâ„¢, that have separately, and together, developed applications to shape the electronic and informational richness of the medical field. The result is that the volume of available data far out distances the capacity of human beings to process it into beneficial evidence. Thus, the goal of natural language processing (NLP) is to transform the wealth of unstructured data into preprocessed or indexed resources fitted to metadata search and document generation [13].



' **Fig. 3** NLP Techniques Used in Healthcare

Various NLP applications do exist in the healthcare domain. The different agencies mostly use NLP for the purpose of analyzing the healthcare reports, with which it identifies the name of the diseases. In healthcare, NER is successfully applied and it is used to find the genetic markers which help in the early prediction of cancer and for finding and isolating the highly infectious patients with extremely resistant organisms. Medical NER is immensely helpful in the generation of structured conceptualization of the texts[13]. With that, dissimilar electronic health records go under comparison which is helpful in the analysis of variability, assessment of care quality, data quality enhancement, and risk profiling. Furthermore, it also improves the design of intelligent healthcare-based systems which can handle unstructured texts.

4. Clinical Applications of NLP in Healthcare

Biomedical NLP has received much attention due to the rapid growth of medical literature, the proliferation of clinical texts, and the significant increase of EMR in recent years. A retrospective study using EMRs generated over 1,000 scientific articles published in 2012 alone. The explosion of data has been accompanied by advances in clinical NLP research, but techniques and systems rarely transfer to improve patient care. This gap may be associated with the focus of clinical NLP research on biomedical problems, such as named entity recognition, and little focus on clinical problems. In the same line, we summarize the clinical applications that have been encouraged by the publication of the SemEval clinical NLP challenge, which is particularly meant to shift the NLP focus to more clinical issues. For a decade the research community has been able to create and test clinical NLP systems by using shared resources, such as annotated data, which are available for ten clinical information extraction and clinical question answering tasks, ending in 2006[13].

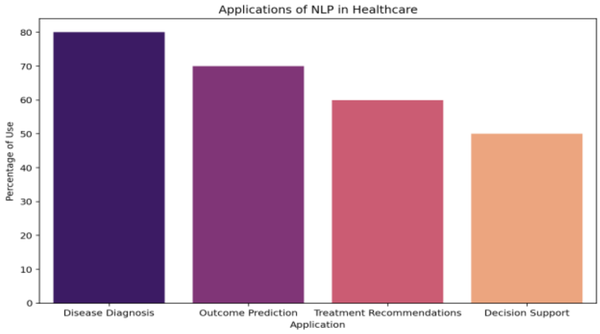


Fig. 3 Applications of NLP in Healthcare

Applications of NLP in healthcare cover three major areas. Clinical applications, like aiding in accurate diagnostics for patient cohorts with complex diseases through conversation, automated empathy for Writing Submission ID trn:oid:::1:2938400128 particular patients, and advising on care pathways. Examples of healthcare applications include quality reporting and assessment, and monitoring continuous care to ensure patient safety and treatment effectiveness. The last layer of applications is concerned with the health system and the public [13]. They ask for expertise in bio surveillance, outbreak prevention, emergency preparation, public health surveillance, medical reporting, patient education, and medical database utilization for finding particular problems within healthcare systems. While NLP plays an important role in these applications, in this chapter we particularly focus on the roles NLP has in terms of clinical applications.

E. EXTRACTING INSIGHTS FROM UNSTRUCTURED MEDICAL DATA

Existing healthcare applications require structured data to produce valuable insights. However, massive volumes of valuable patient information are locked within unstructured data silos. Extracting insights from unstructured data is among the significant challenges in healthcare. The ideal solutions that would solve unstructured data bottlenecks require that (1) clinical information extraction (CIE) solutions are built to work with all kinds of unstructured data, (2) these solutions should be usable for different kinds of information or a combination to increase model trust, (3) the NLP model used should be robust and generalized so that it can work across sources, and (4) ready to be trained incrementally with user feedback, allowing automation to be dictated by a user-dependent training schedule. Due to the complexity of Clinical Notes and free-text information, challenges stem from easy preprocessing (e.g. spelling and grammar mistakes, tokenization, and information normalization), methods that allow domain adaptation, context understanding abilities (e.g. detecting negated terms), overcoming annotation issues (e.g. concept overlap), utilizing community resources (e.g. corpora and dictionaries), working with data variance, scaling, sensitive, and private data, and training deep learning models. Among the available resources, clinical dictionaries and ontologies would help as they are the primary sources of knowledge and metadata in CDMs. They can help increase recall, combine concepts essential for an NLP application, and help in context detection. Due to the complexity of Clinical Notes and free-text information, challenges stem from easy preprocessing (e.g. spelling and grammar mistakes, tokenization, and information normalization), methods that allow domain adaptation, context understanding abilities (e.g. detecting negated terms), overcoming annotation issues (e.g. concept overlap), utilizing community resources (e.g. corpora and dictionaries), working with data variance, scaling, sensitive, and private data, and training deep learning models. Among the available resources, clinical dictionaries and ontologies would help as they are the primary sources of knowledge and metadata in CDMs. They can help increase recall, combine concepts essential for an NLP application, and help in context detection[15,16].

5. CASE STUDIES AND REAL-WORLD IMPLEMENTATIONS

The large pretrained language models have made breaking NLP benchmarks simple. The idea is to take a Transformer-based model with large capacity, train on a large corpus of natural language text, on a variant of the next sentence prediction (NSP) task, and suddenly the model can solve traditional NLP problems. When given a touch of domain-specific adaptation, such as in relation to a particular application area (i.e., specialized fine-tuning), these general-purpose models can also achieve state-of-the-art performance on specific NLP tasks of that domain [16]. The model simultaneously learns the intricacies of human language in the data preparation step, and the task predicts the specificities of that domain in its fine-tuning step. In sum, even small datasets from a specialized domain can yield large performance gains with specialized fine tuning on a pre-trained language model [18]. It does, however, also imply a specific hardware constraint: to include more potential exercises of capabilities or to refer to the optimal model architecture for a specific task in use. This phase has thus in many ways now become somewhat empty because fine-tuning a large pretrained language model will typically result in the best possible model for many applications by any reasonable standard, for example in relation to bandwidth, computational cost, latency, energy consumption, and so on. The core idea of NLP is to teach computers to understand human language. NLP concerns itself with a number of tasks, and in recent times, the human language technology space has seen significant advancement. Identifying sentiments, classifying topics, question-answering, and even language translation have become almost trivial to solve with modern NLP models [18]. This progress has been enabled through two key developments in the field.

II. CONTRIBUTIONS

My contributions in this study is to develop both established and novel natural language processing techniques to extract actionable insights from unstructured medical data, which can have practical implications on clinical detection of charged batteries, fine-grained automated report card generation, and effectively using wearable-device-collected FeNO data for early asthma detection. There are several areas where the current work can be further expanded. Because of the importance of natural language understanding tasks to enable human-computer conversation agents, particularly in healthcare, developing deeper and broader NLP models can further extract more unique and difficult patterns from the text that can benefit a wide range of tasks and applications of healthcare management. Natural language processing (NLP) is a subfield of artificial intelligence (AI) that assists computers in understanding and processing human languages. NLP offers several applications in the healthcare industry by extracting key insights from unstructured data to advance clinical care and operational efficiency. Despite considerable advances in the NLP algorithms and their applications, it has not achieved full maturity. Despite the successes of past research in NLP in the medical domain, its impact and contributions to the aforementioned studies are limited in terms of the system performances that could address the high dimensionality issues with a large number of terms and the high sparsity problem. The uses and applications in other domains such as HelioBase also build upon domain-independent NLP tools specifically used in the domain of cardiology and provide information or advice to both patients and physicians in layman's language. Work presented in other domains needs to be reviewed and performed on the medical datasets, clinical narratives, or the image processing data for the extraction of useful information.

III. SIGNIFICANCE AND BENEFITS

The interest in advancing healthcare practices is responsible for the creation of immense and ever-growing volumes of structured and unstructured non-standardized data that cannot be effectively handled using conventional processes and tools. The majority of the healthcare data is created and used in the form of unstructured notes, dictations, details, discharges, nursing notes, operative reports, pathology reports, and physician progress notes, among others[19]. These sources contain a wealth of information, outside the standard health record, about diagnoses, treatment processes, treatment outcomes, and side effects of treatment encountered. This information has the potential to improve care practices, ensuring efficient use of resources that can be created in an interactive acquisition process responsible for the answer to questions posed during the discovery phase. Such a process should deliver a general overview in detailed probing and allow asking down-to-earth questions with acceptance and methodologies to test generated hypotheses based on the derivative data. In healthcare, two of the most valuable resources, but also the hardest to extract insight from, are time and data. Modern healthcare experiences a data deluge, most of it of the unstructured variety. Despite meaningful advances in natural language processing, from a general computing perspective, the bulk of value gained from mining unstructured information is in systems that process natural language techniques for healthcare. The goal of this chapter is to familiarize the reader with the software used to practice medicine[19]. We begin by introducing natural language processing (NLP) and outline its importance in healthcare. We then continue to review major NLP techniques and present various applications in healthcare, such as data cleaning, clinical coding, adverse drug effects detection, social media tracking, and medical literature mining. We conclude by discussing current "hot topics" in the field and presenting some areas for future research. NLP can manage enormous amounts of unstructured medical texts and turn it into a meaningful and assessable asset, beneficial for new knowledge extraction and information synthesis, at a variety of levels within healthcare systems, including administration and planning health policy, and improving patient care and outcomes. As the use of EHRs containing patient data is a widespread routine, evaluation of the benefits of identifying and addressing unstructured medical texts can be well evaluated through systematic study of necessary, enabling and inhibiting factors[20]. Most of the previous research into the retrieval of medical images has involved non-FDA approved imaging that would not be used in a hospital setting because the researchers would not have access to the proprietary facilities that would make it relevant. Active diagnostic images, MRIs or an X-ray, can be examined by other AI-based tools that have been integrated with the THERESA platform

IV. CONCLUSION

The main objectives were to provide the reader with insight into the core applications of NLP and its intelligent systems and methods in the healthcare sector. The review sought to uncover the prominence of NLP techniques used in the application of critical healthcare domains. It provides an exploration of the latest NLP advancements that improve domain-specific tasks and challenges associated with healthcare data, figure out the effectiveness of state-of-the-art NLP technology and tools, and provide areas for further research. The latest advancements in NLP show that it is a stimulating and challenging field for further research with the potential to address the critical challenges that may arise from the healthcare industry. There are also still some gaps in terms of establishing potent NLP solutions to address some domain-specific requirements. Medical and clinical researchers are in need of convincible NLP-focused universal systems descriptive to communicate technical evaluations effectively. Future research focus must be dedicated towards the development of powerful NLP-oriented common systems, sharing high-quality data, and producing comparable assessments. These challenges can be minimized with ongoing effort and interdisciplinary collaboration. In today's era, there is a continuous explosion of data from various sectors, mainly healthcare. The majority of historical medical and clinical data is recorded in the form of free texts or unstructured data. The extraction and utilization of valuable insights from fields of medicine and healthcare necessitates the application of text data analytics tools such as Natural Language Processing (NLP). Various challenges and issues are associated with the development and performance of NLP tools and techniques.

REFERENCES

1. A. Farzindar, Natural Language Processing for Social Media. Springer Nature, 2015.
2. S. Bandyopadhyay, Emerging Applications of Natural Language Processing: Concepts and New Research. IGI Global, 2012.
3. T. Strzalkowski, Natural Language Information Retrieval. Springer Science & Business Media, 2013.

4. N. Indurkhya and F. Damerau, Handbook of natural language processing. Boca Raton: Taylor & Francis, 2010.
5. E. Reiter and R. Dale, Building natural language generation systems. Casmbridge, U.K. ; New York: Cambridge University Press, 2000.
6. A. Stent and S. Bangalore, Natural language generation in interactive systems. Cambridge: Cambridge University Press, Cop, 2014.
7. E. Krahmer, M. Theune, and Springerlink (Online Service, Empirical Methods in Natural Language Generation : Data-oriented Methods and Empirical Evaluation. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. Ell, User Interfaces to the Web of Data based on Natural Language Generation. KIT Scientific Publishing, 2017.
8. K. Bhavsar, N. Kumar, and P. Dangeti, Natural language processing with Python cookbook. Birmingham: Packt Publishing, 2017.
9. N. Hardeniya, J. Perkins, D. Chopra, N. Joshi, and I. Mathur, Natural Language Processing: Python and NLTK. Packt Publishing Ltd, 2016.
10. J. Perkins, Python text processing with NTLK 2.0 Cookbook : LITE : over 80 practical recipes for using Python’s NLTK suite of libraries to maximize your natural language processing capabilities. Birmingham, England: Packt Publishing, 2011.
11. J. Perkins, Python 3 Text Processing with Nltk 3 Cookbook. CreateSpace, 2014.
12. S. Bird, E. Klein, and E. Loper, Natural Language Processing with Python. “O’Reilly Media, Inc.,” 2009.
13. P. Spyns, Natural Language Processing in Medicine. Leuven University Press, 2000.
14. J. Brender, Medical Informatics Europe ’96 : 13th congress of the European Federation of Medical Informatics : proceedings, Copenhagen, January 1996. Berlin ; New York Etc.: Springer, 1996.
15. K. B. Cohen and D. Demner-Fushman, Biomedical Natural Language Processing. John Benjamins Publishing Company, 2014. Demner-Fushman, K. Cohen, and B. Webber, Introduction to Biomedical Natural Language Processing. 2013.
16. R. Satapathy, E. Cambria, and A. Hussain, Sentiment analysis in the bio-medical domain : techniques, tools, and applications. Cham: Springer, 2017.
17. A. Kao and S. R. Poteet, Natural language processing and text mining. London Springer, 2007.
18. P. Jackson and I. Moulinier, Natural language processing for online applications : text retrieval, extraction and categorization. Amsterdam ; Philadelphia: John Benjamins Pub, 2007.