International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

IJISAE

ISSN:2147-6799 www.ijisae.org Original Research Paper

Leveraging Natural Language Processing in Electronic Health Records for Enhanced Healthcare Decision-Making: A Systematic Review

Swetha V Padmavathi Polisetty[1], G.Deepthi[2]

Submitted:13/03/2024 **Revised**: 28/04/2024 **Accepted**: 05/05/2024

Abstract: These Natural language processing (NLP) is frequently used in Electronic Health Records (EHRs) to extract clinical insights; nevertheless, issues with automated tools, annotated data, and other constraints prevent NLP from being fully exploited for EHRs. To comprehend these limitations and investigate novel prospects, this research compares and analyzes different Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP) methodologies. The work covers seven main areas: classifying medical notes, recognizing clinical terms, summarizing texts, developing advanced AI models, extracting important information, translating medical language, and applying NLP to various other healthcare tasks— We looked through 261 articles from 11 databases in a systematic review and selected 127 of them for a detailed reading. Three novel goals are combined in this study: a handwritten prescription interpretation system based on GPT-3, multilingual clinical note extraction using BERT variations, and PRISMA-compliant label extraction from radiology reports using Vision Transformers. The findings show that the majority of data in electronic health records are unstructured, and that the most popular uses of machine learning and deep learning approaches are prediction and categorization. ICD-9 categorization, clinical note analysis, and named entity identification are some of the major use cases. The results of our suggested systems were encouraging: the BERT variations handled multilingual clinical notes efficiently, the GPT-3 system transcribed handwritten prescriptions properly, and the Vision Transformers increased the radiological report label extraction efficiency.

Keywords: *Medical NLP, Machine Learning, Electronic Health Records, Artificial Intelligence*

1. Introduction

The process of gaining valuable insights from massive volumes of unstructured data has become more difficult and more opportunity-driven with the digitalization of healthcare records. Three cutting-edge initiatives that optimize clinical workflows and improve patient care via the use of cutting-edge AI techniques are the subject of this study, which attempts to solve some of these issues.

Electronic health records, or EHRs are computerized summaries of medical procedures and evaluations that are becoming more and more common and crucial for research, management, and delivery of healthcare [1]. Both organized and unstructured data may be found in EHRs [2]. Conversely, clinical notes and discharge summaries which are created by medical staff are examples of unstructured data. Clinical records or worldwide, the use of EHRs has grown significantly.

The initial goal is to create a cutting-edge patient assistance system with an automated reader designed to correctly analyze handwritten prescriptions using GPT-3. Because doctor's handwriting varies so much and their handwriting is sometimes unreadable, handwritten prescriptions are a frequent source of mistakes in healthcare. This system seeks to efficiently transcribe and interpret handwritten prescriptions, minimizing drug mistakes and enhancing patient safety by utilizing GPT-3's potent language comprehension skills

The second goal is to employ BERT Variations to extract multilingual information from clinical notes. Although clinical notes include a wealth of important information, reliable data extraction and analysis can be difficult since they are frequently written in many languages. This research aims to improve the extraction of important medical information from clinical notes, independent of the language spoken, by utilizing multilingual BERT models. This will increase the usefulness and accessibility of medical records in variety of healthcare settings.

The third goal is to use Vision Transformers to automatically extract picture labels from radiological reports with the highest level of accuracy and efficiency. Critical diagnostic data included in radiology reports must be appropriately labeled in order to support efficient treatment planning and medical research. With its cutting-edge image processing capabilities, Vision Transformers provide a viable option for automating this procedure. In order to improve the effectiveness and precision of radiological evaluations, this project will create a system that will automatically extract and identify information from radiology pictures

The volume of information that may now be referred to as

[1] *School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh – 522237, India*
ORCID: 0009-0008-2909-1336
School of Computer Science and Engineering, VIT- AP University, Amaravati, Andhra Pradesh – 522237, India
ORCID: 0000-0003-0712-6899
** Corresponding Author Email: deepthi.g@vitap.ac.in*

"Big Data" due to the growing use of EHRs includes the modification and application of significant amounts of data that have been gathered in EHRs. We need to create computer tools to organize, evaluate, and find patterns in data because human brains can't handle the complexity of analyzing and understanding large amounts of data. The next stage involves turning all of this massive healthcare data into knowledge by applying techniques like data mining and natural language processing, which are crucial parts of data analytics on big data from EHRs and support the growth of an EHR ecosystem.

One form of ML integrated technology that has recently gained popularity is autoML [13]. Its purpose is to increase the application of machine learning (ML) algorithms and make their deployment easier across many sectors, including the healthcare sector [14]. While it is still a relatively new technology, autoML has already found applications in translational medicine, bioinformatics, diabetes, Alzheimer's disease, electronic health record (EHR) analysis, and medical imaging [15]. It hasn't, however, been thoroughly examined in terms of its potential applications for processing clinical notes, a crucial EHR component. Additionally, ensuring the security of patient data is a critical consideration when deploying automated systems enabled by machine learning. However, no research has thoroughly examined the techniques that may be used to ensure the privacy of medical data or identified challenges related to patient data protection.

Apart from this, a number of ways have been devised to enable EHRs to manage clinical duties, yet, the distinct vocabulary and clinical idioms employed by doctors continue to provide difficulties for health information research [16], [17]. In theory, these methods are still in the early stages of development, and it will make some time before they can choose a precise and accurate model for practical use. This brings up the biggest issue facing NLP researchers today, using computer technology to interpret medical text data and make decisions. To help NLP be used effectively in modern healthcare, new classification schemes are needed. Primarily, the goal of this project is to address the deficiencies that have been found in EHRs-NLP applications for healthcare and develop efficient techniques for EHR analysis that will benefit the research community.

These are the goals that our review aims to achieve. Initially, our goal is to examine the NLP method in EHRs, paying particular attention to several cutting-edge models. We then go over the DL and ML paradigms that are applied to the analysis of EHRs, namely clinical free text. Third we pinpoint the main difficulties in classifying clinical notes. Finally, we look at how academics have applied their models clinical note management in the healthcare sector.

When taken as a whole, these initiatives offer a through strategy for utilizing cutting-edge ML and NLP methods to tackle major healthcare data management issues. This work intends in order to significantly advance the field of health informatics by automating picture label extraction, strengthening multilingual information extraction, and improving the understanding of handwritten prescriptions.

2. Related Work

2.1. Selection Criteria

A. Inclusion Criteria: The following topics have been covered in the literature: ICD-9 multi-label classification, clinical entity recognition, automatic clinical narrative summarization, medical concept embedding, healthcare dialog systems. Only complete conference papers or peer-reviewed journal articles were included. Studies that only employed models or frameworks based on ML or DL for EHR analysis were eligible for inclusion. Research must also have been concentrated on using ML and DL techniques to analyze and identify clinical narratives.

B. Exclusion Criteria: Excluded from consideration were research projects that were released as preprints, included preliminary work, or lacked peer assessment. Review papers and editorials were also excluded from consideration. Following the first screening, the quality of the articles that were retrieved for full-text analysis was also assessed.

2.2. Electronic Health Record

Many healthcare facilities have recently switched from using traditional paper-based medical report records to Electronic Health Capture (EHR) systems, which capture patient's longitudinal medical information in an electronic format on a data repository. Like other types of electronic health (e-health) infrastructure, EHRs use information and communication technology to digitize and automate processes associated with the provision of healthcare. Due to this swift transformation, EHR systems, now have access to a vast volume of clinical records, many of which include valuable patient data. Beyond the obvious space savings associated with moving from paper to digital records, EHR has been essential in enhancing decision-making, promoting healthcare efficacy, and enhancing management.

The electronic health record (EHR) system is linked to medical decision support systems (DSS), which assist physicians, staff members, and management in making choices. Among other things, it makes it easier to make prompt and accurate decisions on data analysis, diagnosis, and invoicing. Making Quick, accurate decisions on diagnosis, lab testing, processing invoices and payments, and data analysis is made simpler by it. We show an example of an e-health infrastructure for the provision of healthcare. A number of e-health infrastructure components were made available, such as an electronic health record, a knowledge base, an EHR and a physician support system.

2.3. Natural Language processing

A branch of study called natural language processing (NLP) was born out of the collaboration of linguistics and computer science. Its goal is to enable computers to comprehend and interpret spoken and written human language. Text classification, machine translation, summarization, and Named Entity Recognition (NER) are a few typical NLP tasks. There are several NLP system kinds. NLP can be divided into three main groups: statistical NLP, neural NLP, and symbolic NLP.

2.3.1. Symbolic NLP

This method relies on a predefined set of grammatical rules intentionally coded by linguists. It was the first form of NLP to emerge and is still in use today. This branch of NLP involves working with regular expressions, which are patterns used to search for and match specific characters or sequences within a text. This branch of NLP deals with regular expressions, which are search patterns used to find and match one or more characters in a string. Establishing a system of rules takes effort and specialized knowledge. This stage requires careful consideration and effort because it is essential to the functionality of these algorithms. These systems ability to develop with little to no data requirements is one of their advantages. The purpose of rule-based systems goes beyond only word recognition in text.

2.3.2. Statistical NLP

Many times, a collection of rigidly defined rules cannot adequately convey the intricacy, subjectivity, and subtleties of a whole language. Systems that rely on symbolic NLP fail in these situations.

The primary issue with statistical methods is that they only consider an important measure for each word, which ignores semantic similarity between words. Therefore, terms like “thorax” and “chest” will be handled as if they are entirely unrelated. Furthermore, the context in which a word is used is not considered by these algorithms. Homonyms, like words with the same spelling but different meanings, are considered identical when representing text as a Bag-of-Words (BoW). BoW is a fixed-length vector where each position corresponds to a word in the vocabulary, and the value at each position indicates how often that word appears in the document. Additionally, sequences of n words, called n-grams, can also convey the same information.

Instead of using single words (unigrams), we can use pairs or triplets of words (bi- or tri-grams) to give more context. For example, the word "chest" can be better understood when paired with other words like "treasure chest" or "chest X-ray." Another option is to use one-hot encoded vectors to represent words. These vectors are the same size as the vocabulary, with only one value set to one and the rest set to zero. However, neither of these methods captures the similarity or difference between words. Additionally, using one-hot encoded vectors can be challenging because they create very sparse representation of data. Another popular option is statistical language models, also referred to as language modeling [8]. These models help in understanding the context of words and phrases by considering the probability distribution across a sequence of words. However, these models face a challenge known as the "curse of dimensionality," similar to one-hot encoded vectors. As a result, statistical language models often rely on n-grams, which assume that the context of a word depends only on the n words around it, rather than the entire set of words in the vicinity. While this approach reduces the problem of sparse findings, it doesn't completely resolve the issue. Therefore, there's still a need for better and more concise representations of words.

2.3.3. Neural NLP

Convolutional Neural Networks (CNNs) are commonly used for processing images, but they've also been adapted for processing text. In tasks like sentiment analysis or identifying spam, CNNs can be useful. Instead of image data, they work with matrix word embeddings. CNNs have two choices: they can either learn their own embeddings or use pre-trained embeddings, like those from a word2vec model. Pre-trained embeddings are good at understanding the meaning of words, as they're trained on huge datasets. However, CNNs might not be ideal in some cases because they can lose information about the order of words during convolution and pooling processes.

The meaning of statement can be significantly impacted by the word order. The Recurrent Neural Network (RNN) is a novel kind of neural network that was created as a result of the sequential structure of text and the natural requirement to take word order into account in particular applications. Because these neural networks analyze text sequentially, they are able to record the relationships between words that appear in a specific order. Put differently, RNNS identify patterns throughout time, where CNNs identify patterns across space. Regarding the length of input data, RNNs have an additional benefit over CNNs: they can model text of varying lengths from single phrases to complete texts. As a result, the network may take into consideration distant semantic relationships.

RNNs consist of several discrete building components. They may be viewed as several, chain-like implementations of a very basic feed-forward artificial neural network (ANN), each of which passes information on to its successor and has identical weights and activation functions. This design can take into consideration the dependencies between the components in a sequence rather than handling each input item separately. Owing to this attribute, these networks are referred to as having memory.

3. Methodology for Patient Aid System for Hand Written Prescriptions

Even though most studies have focused on recognizing human handwriting, they still haven't come up with a practical solution because of the significant variations in individuals' writing styles. In the medical field, doctor handwriting recognition has emerged as a serious problem in the contemporary world. The use of Latin acronyms and the physician's unreadable cursive handwriting make it difficult to identify the prescriptions that they write. Achieving 99.99% accuracy in recognizing handwritten cursive words solely through OCR algorithms has proven challenging. As a result, researchers have explored combining machine learning concepts with neural networks in attempts to achieve this goal.

Ultimately, the procedure aims to turn handwritten or printed text into an image by utilizing the principle of Optical Character Recognition (OCR). In order to verify that the outcome matches the expectation, the previously processed photos are subjected to a pre-trained model, the information extraction process should be carried out in two stages: spelling correction and classification. The training model, created using the NLP NER method, will utilize the unorganized text data obtained from scanned prescriptions. Because of the specific domain, we used a modified NER model for this purpose.

After the text data has been classified by the model, it is important to verify that the medicine name has been accurately classified. This is because the medicine name has to be 100% accurate in order for the application to function properly. If it’s not correct, there ought to be a way to make it more accurate, spell correction mechanisms can be used to increase accuracy by attempting to discover ideas through character substitution and interchange, as well as by adding new characters to existing words. The best word from the list of recommendations will then be determined by looking the word with the least amount of edit distance.

With the name of the retrieved medication, the allergy potential category can then be determined. Using the allergy category data set, it is necessary to determine whether the specific drug name has been recorded at any allergy potential category in order to determine the allergy possible category.

This method is divided in two parts: the scanning of blood reports and prescriptions. Prescription scanning begins with transferring the scanned medical prescriptions as pictures to the systems, where the handwritten content is machine-encoded. The system extracts the medication information by categorizing the machine-encoded texts, including the name, dose, and directions, and then stores the information in the database as prescription details. The relevant data, which includes indications, brand names, states, and overdose symptoms, is taken out of the drug repository and mapped into the main database.

The uploaded blood report image has to be read in order for OCR to turn it into a text document. By scanning patient blood reports via the mobile application, this technology is able to evaluate them.

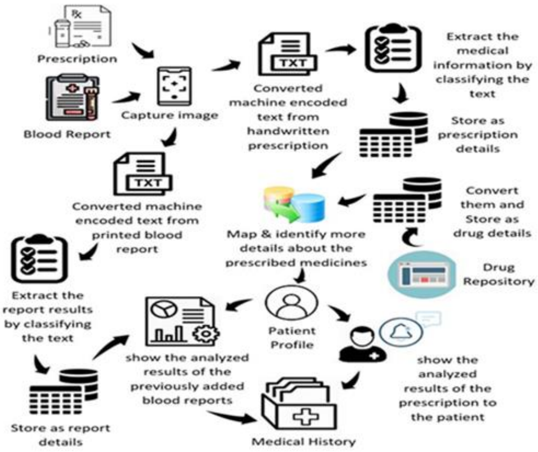

Fig 1. Healthcare Data management from Prescription to patient profile

Fig 2. Segmented words from prescription

Dilation is a process that enlarges the pixels in an image and adds pixels to the edges of objects. It works by setting a pixel to 1 if any of its neighboring pixels have the same value, and the output pixel value is the highest value among all the pixels in the neighborhood. This effectively expands the object in front of the camera or increases the amount of white space. Contour analysis involves identifying the relevant contour lines around the image, forming a rectangular box to better identify the region. The identified regions are then kept as separate images.

4. Multilingual Information Extraction using BERT Variants

Information in Multiple languages using BERT variations for extraction from clinical notes is a complex method to enhance the comprehension and processing of medical information in different languages. The language variety of clinical notes presents substantial hurdles for automated systems since they frequently contain information that is essential for patient care, research, and healthcare decision-making. This approach makes use of BERT's sophisticated contextual knowledge to handle the complexity and diversity of clinical language by utilizing language-specific versions such as multilingual BERT (mBERT). Preprocessing is the first step in the process, during which clinical notes are cleaned, standardized, and extraneous material is removed. Medical terminology and abbreviations are then normalized. Then, to better meet the unique requirements of clinical information extraction and the subtleties of medical language, multilingual BERT models are adjusted on a dedicated corpus of clinical notes pertinent entities and relationships-such as patient names, prescriptions, symptoms, diagnoses and treatment plans are found and extracted. To guarantee accuracy and relevance, the retrieved data is subjected to post processing and validation. This may involve human evaluation by medical specialists or cross-referencing with organized medical databases. This method improves patient care and streamlines clinical processes by enabling health care systems to automatically extract and use vital medical information from multilingual texts. It provides multilingual medical text processing with a more unified and effective solution by decreasing reliance on language specific models.

In this work, we offer an end-to-end system that creates a biological knowledge network from clinical textual data that is unstructured and hence difficult to evaluate using a variation of BERT models. Medical choices will be based on the findings of this knowledge graph's development and analysis. This would help medical practitioners examine each entity independently as well as all the connections among them. Additionally, it would make it simple for doctors to identify the connections between a medication and clinical notes, enabling close monitoring of these medications.

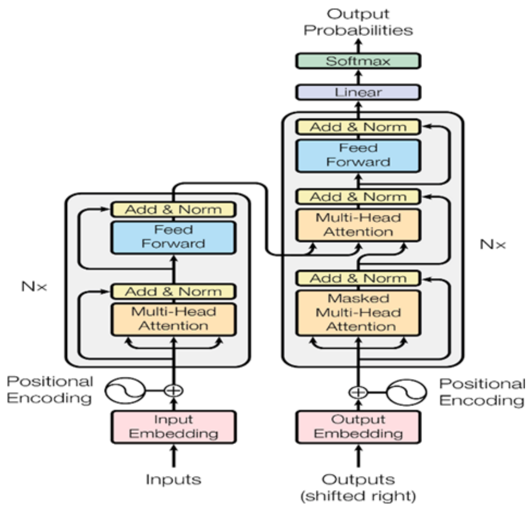

Fig 3. Transformer Architecture

4.1. BERT

Bidirectional Encoder Representations from Transformers, or BERT, is a ground breaking approach in natural language processing (NLP) that was created by Google researchers. BERT processes text bidirectionally, which means it concurrently takes into account the context from both sides, in contrast to standard models that process text either from left to right or from right to left. As a result, BERT is better able to comprehend linguistic subtleties and accurately convey the meaning of words in context.

BERT is based on the Transformer Architecture, which used a process known as self-attention to estimate the relative value of words in a phrase. BERT can perform well on a variety of NLP tasks, including named entity identification, text categorization, and question answering thanks to this design. BERT gains a profound grasp of language patterns by pre-training on a large volume of text data. This allows BERT to learn how to predict masked words in a phrase and evaluate whether two given sentences are successive sections of the text.

The status of NLP has greatly evolved since BERT's launch, allowing for improved accuracy and more reliable performance in a range of applications. It differs from earlier models in that it can comprehend context in both directions, which makes it an effective tool for developers and academics working on language-related projects.

4.2. BERT Variants

1. DistilBERT: DistilBERT [33] is a more compact, quicker, less expensive, and lighter variant of BERT. Despite only having half of BERT's parameters during training, DistilBERT performs 97% as well as BERT. It is challenging to handle the 110 parameters in BERT-base and the 340 parameters in BERT-large. The distillation process is used to shrink these massive models to tackle this issue. Distil-BERT's overall design is the same as BERT's overall design is the same as BERT's with the exception of the removal of the pooler and token embeddings. This leads to a halving of the number of layers, which greatly enhances computing efficiency.

2. BioBERT: A BERT variation pretrained on a biological dataset is called BioBERT [24]. Pretraining involved taking the usual BERT model's weights and pretraining it on medical datasets such as PubMed abstracts and PMC.

3. BioClinicalBERT: Beyond organized data, such test results and prescriptions information, clinical notes include personal information about their patients. However, due to their high dimensionality and sparsity, clinical notes have not been used as much as organized data. Clinical BERT variation with a focus on clinical notes is called BioClinicalBERT [3]. It draws attention to excellent connections between medical ideas as determined by human judgment.

4. Bio-discharge-summary: A biomedical specific version

of BERT called Bio-Discharge-summary [3] was trained using clinical notes from the MIMIC hospital and initialized using BioBERT. After then, it was improved with just minor architectural changes for the five jobs.

4.3. Knowledge Graph Construction

1. Named Entity Recognition (NER) is a part of information retrieval that focuses on identifying and classifying specific entities mentioned in text. These entities can include names of people, places, organizations, medical codes, time expressions, quantities, and more. The goal is to categorize them into predefined groups. The main categories include names of people, places, and institutions (like organizations); time expressions; and numerical values like currency and percentages. These categories can be expanded to suit specific needs in different applications.

2. Graph Analysis: We created a step-by-step process to transform unstructured clinical notes into a format that's easier to understand. First, we prepare the clinical note by addressing any coreference issues. Then, we use a model called BioclinicalBERT+CRF to predict the entities mentioned in the note. Next, we preprocess the note again to prepare it for another model called BioClinicalBERT+softmax, which helps extract connections between the predicted entities. A table is created to link each medicine to all of the things that are connected to it after the links between the entities are predicted (Fig. 4).

3. Conference Resolution: Finding and linking assertions in a text that pertain to the same real-world item is known as conference resolution (CR), and it is one of the main goals of natural language processing (NLP). To eliminate ambiguity and improve text coherence, coreferents-a grouping of these expressions are the aim of CR. For many sophisticated NLP applications, including summarization, question answering, and text comprehension, this procedure is essential.

Coreference is the word used to describe the phenomenon where in many phrases in a text refer to the same thing. Pronouns, proper names, and descriptive phrases are examples of these expressions. For example, take the following sentences: “Alice visited the market”. The pronoun “she” refers to “Alice” who purchased some apples. The fundamental step in resolving coreferences is identifying “Alice” and “she” as coreferents.

Pronominal coreference nominal coreference, and proper name coreference are among the several forms of coreference. Pronouns that refer back to nouns or noun phrases sated previously in the text are known as pronominal coreference. An example of this would be “John has a car.” “Every day he drives it,” in which “he” and “it” stand for john and an automobile, respectively. When two noun phrases such as “The president” and “The leader” refer to the same individual. The leader stressed. When two proper names “Barack Obama” and “president Obama” refer to the same thing in a text, this is known as proper name coreference.

4. Relation Extraction: In NLP, relation extraction (RE) is an essential activity that aims to detect and categorize semantic links between items in a text. This work entails finding instances-people, locations, organizations, or other things where two or more of these items are related in a certain way. These relationships are then classified into predetermined semantic categories. In the line “John Doe works at OpenAI” for instance, “OpenAI” is an organization entity, while “John Doe” is a person entity. We can categorize the relationship between these businesses as “employed by”. These connections offer useful structured data that may be applied to a number of different tasks, such as question answering systems, knowledge base populating, and information retrieval.

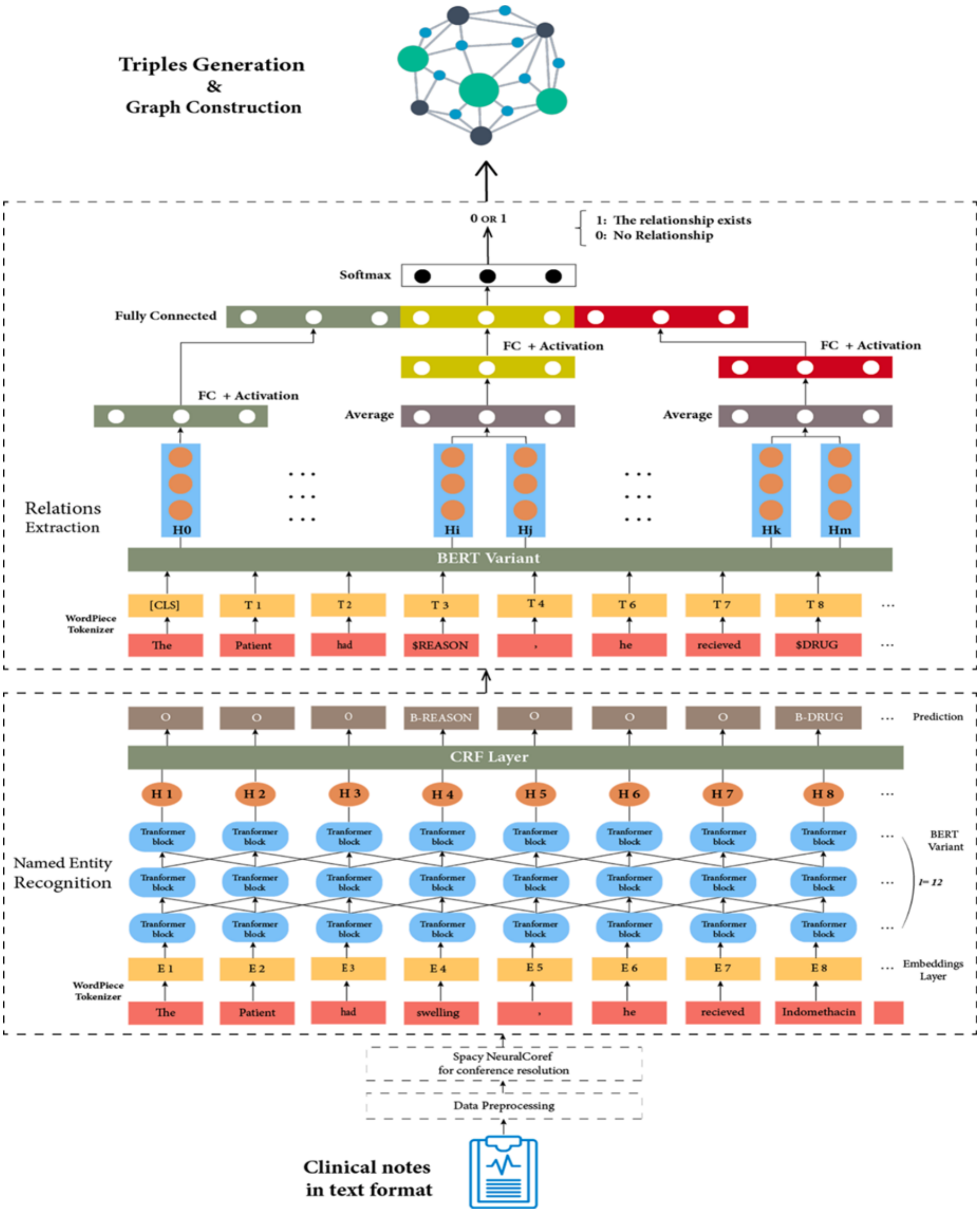

Fig 4. Information Extraction from Clinical notes using variants of BERT

5. Multilingual information Extraction using BERT variants

5.1. Radiology

Radiology is a branch of medicine that uses various imaging techniques to examine, diagnose, and treat diseases. There are two main types: Interventional radiology, which assists in guiding treatments such as inserting catheters or performing biopsies, and diagnostic radiology, which involves examining a patient's organs to detect any abnormalities. A licensed physician with expertise in the interpretation of medical pictures known as a radiologist.

A report is created by a radiologist after evaluating a radiological exam. We will now go into depth about the format and substance of radiology reports.

5.2. Radiology Reports

A radiologist's description and interpretation of the findings of a medical imaging examination are recorded in radiology reports. Hospitals often save reports and the related radiological pictures on the PACS. It is essential to comprehend their vocabulary and structure in order to design NLP tools that are suitable for processing this particular kind of data. An expert reanalysis of the test or report can yield ground truth labels for automated labeler training.

Reports are often organized into parts, however different hospitals may use different formats. In addition to the patient's name, gender, and age, a radiology report typically includes the date of the examination and the name of the radiologist who interpreted the results.

The clinical history section of a medical report provides symptoms, and any pertinent prior medical problems (if any).

• Technique: Includes exam-specific technical information. For instance, it might specify the kind of test and/or whether or not contrast was employed.

• Comparison: Includes a brief comparison between the test being taken now and any previous exams that have been taken.

• Results: outlines the radiologist's observations in the various bodily areas examined. This section often focuses on indicating whether or not specific clinically significant parts of the picture are normal. These characteristics can vary depending on the patient, organs, and imaging modalities.

• A description of the findings and, if feasible, a diagnosis based on the facts at hand are included in the impression.

5.3 Health Concept Representation Systems

Radiology report portions are written as free text, which makes indexing and organizing them challenging. Because universal health concept representation methods lessen the ambiguity inherent in natural language, they considerably aid in the organization of knowledge and the extraction of information from these documents. There are three primary types of systems used in information organization: ontologies, controlled vocabularies/thesauri, and terminologies. Terminologies consist of sets of terms related to specific topics or fields, while controlled vocabularies or thesauri establish connections between terms, illustrating their relationships. Ontologies, on the other hand, represent more complex systems, organizing concepts and their interrelations in a structured manner. Since ontologies are intended for computational usage, logical definitions of the concepts are included as well. There are idea representation systems for both general medicine and radiology.

Creating NLP applications for radiology reports can significantly benefit from using terminological resources. Essentially, these resources serve as templates for generating accurate labels for images, establishing patterns that are considered the truth. Terms from these resources can be directly marked or annotated to provide valuable information, this would result in an extremely large number of label categories, or they can be annotated using a pre-established pooling system, which would result in a smaller set of labels by translating patterns of terms into a smaller set of labels. To put it briefly, consistent and universal labels are ensured by ground truth annotations that are completed utilizing terminological resources as references.

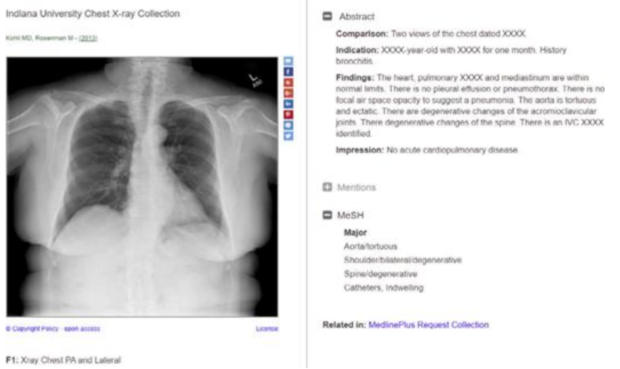

Fig 5. Example of a data instance from the Indiana data set, as presented in the Openl repository

5.4 Radiology images and their annotation

In radiology, different imaging techniques are used to examine patients. These include computed tomography (CT) scans, computed radiography (X-ray), magnetic resonance imaging (MRI), and ultrasound scans. Every modality has pros and cons and it's appropriate for usage in certain situations and applications. When an imaging exam is conducted, a radiology report is usually generated, regardless of the modality.

Radiologists analyze and annotate medical imaging pictures to create hand-labeled data sets for supervised machine-learning algorithms. Through this approach, ground truth picture labels are produced, which machine learning models may use to identify radiological abnormalities or illnesses. Nonetheless, the radiology reports that go with the pictures already provide unstructured explanations and analyses of the pictures Images and written reports from the Indiana Network for Patient Care are included in this data collection, for patient care are included in this data collection, together with ground truth labels that have been carefully annotated using 76 RadLex and 101 MeSH vocabulary words by skilled human annotators. After automatically anonymizing the reports and images, the outcomes were manually confirmed.

6. Discussion

The effects of iatrogenic injuries make patient safety a crucial healthcare issue. In critical care, medication mistakes are common, harmful, and avoidable. The implementation of the EHR system has increased the number of medical mistakes, even though the main goal of the system's adoption was to address various issues with medical healthcare delivery. Nevertheless, human factor research conducted in non-medical contexts indicates that mandating more attentiveness for medical professionals would not result in a significant increase in safety. Redesigning malfunctioning systems and identifying problems seem to be more efficient ways to lower human error.

Complex natural language processing (NLP) models are typically integrated into electronic health record (EHR) systems, where they serve as complete meaning suggesters during information retrieval and auto-completion tools for drafting medical reports. Nonetheless, the model carries out these tasks while keeping in mind the context of the acronym's use in the report statement. Several strategies have been put out over time. Notably, the consensus has been that artificial intelligence-based solutions are the most effective way to find a comprehensive form for acronyms in medical notes. The many AI approaches that have been used throughout the years include machine learning-based techniques, statistical approaches, and knowledge-based methods. The majority of the work currently in publication has the notable drawback of using datasets consisting only of hundreds or thousands of texts to train machine learning models.

Inadequate design and training of the algorithms may result in biased outcomes that adversely affect particular patient populations, such as those with particular medical problems or demographic traits. This might worsen already existing health inequities or lead to biased treatment. Concerns exist around the openness and accountability of deep learning and natural language processing algorithms. In addition to making sure that choices are reasonable and fair, healthcare practitioners need to be able to explain how these algorithms arrive at their conclusions. It should be possible for patients and other interested parties to comprehend and challenge the judgments that these algorithms make.

Statements of denial and ambiguity, for instance, might constrain how a particular discovery is interpreted. NLP systems require consideration of these modifier expressions to provide appropriate picture labeling. Additional characteristics of mentions include location pattern and intensity. As a result, radiology reports provide a difficult combination of data that must be used as effectively as feasible.

Apart from technological aspects, compliance with healthcare regulations and standards become a critical priority. This section examines the complex issues involved in making sure AI applications in healthcare comply with current laws while accounting for regional differences in compliance specifications. Furthermore, there is a serious risk of relying too much on AI-driven solutions, which raises concerns about a reduction in human oversight and decision-making. It becomes crucial to strike a careful balance between maintaining the human touch in healthcare and utilizing technology to increase efficiency to guarantee that patients receive individualized, compassionate treatment. Furthermore, the deployment of sophisticated AI systems especially LLMs has a substantial financial impact. It is important to carefully assess the expenses involved in obtaining, adopting, and maintaining these technologies, especially for healthcare organizations that have restricted funding. These intricate problems highlight how delicately integrating AI into medical procedures.

7. Conclusion

The primary challenge of the research has been identified as the difficulty in identifying a doctor's handwriting, and it is a well-known truth that the majority of systems that have already been put into place have also failed to come up with a workable answer for this. Additionally, it enables users to create and manage their patient profiles, in which they may save the recorded prescriptions and access comprehensive drug descriptions, potential allergies, and adverse effects.

Several strategies, including data augmentation, adding more CNN layers, increasing input sizes, and adhering to the vanilla beam search decoding algorithm, can be used to increase the accuracy level of this methodology. If further work is needed to expand the dataset as a result of this research, it may be done so effectively by combining a newly created customized dataset with the same model. As a bonus to this research, blood reports in any suitable category or format can be used in place of the use of blood reports.

In this study, we provide a comprehensive and accurate process that can be used to build a network of biomedical information derived from various clinical data sources, offering valuable insights a valuable starting point for analysis. Keep in mind that the abundance of patient data accessible is a barrier when using medical textual data. Additionally, many of the activities performed by medical professionals are time-consuming, repetitive, and involve searching for information. This requirement has prompted researchers to look for alternative, quicker, and more creative methods, accessing information through a system that's organized like a knowledge graph. Most studies in the literature focus on specific parts of building a knowledge graph, but they often don't combine these aspects to create the entire graph. Sometimes, they also overlook the analysis phase, which is crucial for understanding the significance of constructing the graph.

From a medical and patient standpoint, our goal is to create a search engine that will allow them to make use of the features of this knowledge network and search for the information they require contextually. However, our goal is to expand this knowledge graph's application area so that it may be utilized in other industries, including the pharmaceutical industry. Annotated medical image data sets may be produced with great ease if it is possible to use picture datasets with automated labeling natural language processing (NLP) on the data found in radiology reports. This piece tackles the subject in a few different ways: First, a brief introduction to radiology reports and NLP is given. Following that, a review of pertinent literature is conducted. The process of automatically extracting labels from radiology reports has evolved from solely using rule-based methods to incorporating machine learning techniques. Often, a combination of these approaches is utilized for specific tasks within a larger process. While it may not be explicitly stated, most current research leans towards employing neural networks over rule-based algorithms, which are relatively new in this field. Although direct comparisons are challenging, most advanced label extraction methods perform well, with the distinction between rule-based and machine learning techniques being relatively minor.

Research on physiological illnesses, including dementia and geriatric mental health, has been highlighted as a potential field. A variety of models and techniques for feature extraction appropriate for these tasks are being investigated in this continuing study. Despite the deep learning algorithm's remarkable success in natural language processing, we discover that applying them to the biological domain is still challenging. DL models have several drawbacks in comparison to traditional ML models, which are commonly applied to health records. These drawbacks include data

availability, interpretability issues, and the challenge of handling domain-specific textual data. Notable is the fact that hundreds of workstations and costly GPUs are needed for DL-based algorithms to perform better than other approaches.

Conventional approaches are currently dominant, while cutting-edge NLP techniques, including transformer-based models for free text analysis, have not yet been widely employed, As a result, we are curious as to whether transformer-based methods will eventually take precedence in clinical NLP.

References

[1] M.R. Cowie, J.I. Blomster, L.H. Curtis, S. Duclaux, I.Ford, F.Fritz, S. Goldman, S. Janmohamed, J.Kreuzer, M. Leenay, et al., Electronic Health records to facilitate clinical research, Clin. Res. Cardiol. 106 (1) (2017) 1-9, http://dx.doi/10.1007/s00392-016-1025-6.

[2] H. Consultant, Why unstructured data holds the key to intelligent healthcare systems [Internet], 2015, Atlanta (GA): HIT Consultant URL https://hitconsultant.net/2015/03/31/tapping-unstructured-data-healthcares biggest-hurdle-realized/.

[3] J. Liang, Y. Li, Z. Zhang, D. Shen, J. Xu, X. Zheng, T. Wang, B. Tang, J. Lei, J. Zhang, Adoption of electronic health records (EHRs) in China during the past 10 years: Consecutive survey data analysis and comparison of Sino-American challenges and experiences, J. Med. Internet Res. 23 (2) (2021) e24813–e,http://dx.doi.org/10.2196/24813.

A. Hodgkins, J. Mullan, D. Mayne, C. Boyages, A. Bonney, Australian general practitioners' attitudes to the extraction of research data from electronic health records, Aust. J. Gen. Prac. 49 (3) (2020) 145–150, http://dx.doi.org/10.31128/ AJGP-07-19-5024.

[4] K. Cairns, M. Rawlins, S. Unwin, F. Doukas, R. Burke, E. Tong, A. Henderson, A.C. Cheng, Building on antimicrobial stewardship programs through integration with electronic medical records: The

[5] Australian experience, Infect. Dis. Ther. 10 (1) (2021) 61–73, http://dx.doi.org/10.1007/s40121-020-00392-5.

[6] U. Naseem, M. Khushi, S.K. Khan, K. Shaukat, M.A. Moni, A comparative analysis of active learning for biomedical text mining, Appl. Syst. Innov. 4(1) (2021) 23, http://dx.doi.org/10.3390/asi4010023.

[7] Y.H. Bhosale, K.S. Patnaik, Application of deep learning techniques in diagnosis of COVID-19 (Coronavirus): A systematic review, Neural Process. Lett. (2022) 1–53, URL https://link.springer.com/article/10.1007/s11063-022-11023-0.

[8] Y.H. Bhosale, K.S. Patnaik, IoT deployable lightweight deep learning application for COVID-19 detection with lung diseases using RaspberryPi, in: 2022 International Conference on IoT and Blockchain Technology, ICIBT, IEEE, 2022, pp. 1–6, URL https://ieeexplore.ieee.org/document/9807725.

[9] A.L. Beam, I.S. Kohane, Big data and machine learning in health care, JAMA 319 (13) (2018) 1317–1318, URL https://jamanetwork.com/journals/jama/article-abstract/2675024.

[10] Y.H. Bhosale, S. Zanwar, Z. Ahmed, M. Nakrani, D. Bhuyar, U. Shinde, Deep convolutional neural network based COVID-19 classification from radiology XRay images for IoT enabled devices, in: 2022 8th International Conference on Advanced Computing and Communication Systems, Vol. 1, ICACCS, IEEE, 2022, pp. 1398–1402, URL https://ieeexplore.ieee.org/document/9785113.

[11] A.F. Leite, K.d.F. Vasconcelos, H. Willems, R. Jacobs, Radiomics and machine learning in oral healthcare, PROTEOMICS–Clin. Appl. 14 (3) (2020) 1900040, URL https://onlinelibrary.wiley.com/doi/full/10.1002/prca.201900040.

[12] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, J. Dean, A guide to deep learning in healthcare, Nat. Med. 25 (1) (2019) 24–29, http://dx.doi.org/10.1038/s41591-018-0316-z.

[13] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, F. Hutter, Efficient and robust automated machine learning, Adv. Neural Inf.

Process. Syst. 28 (2015) URL https://proceedings.neurips.cc/paper/2015/file/ 11d0e6287202fced83f79975ec59a3a6-Paper.pdf.

[14] J. Waring, C. Lindvall, R. Umeton, Automated machine learning: Review of the state-of-the-art and opportunities for healthcare, Artif. Intell. Med. 104 (2020) 101822, http://dx.doi.org/10.1016/j.artmed.2020.101822.

[15] A.A. Borkowski, C.P. Wilson, S.A. Borkowski, L.B. Thomas, L.A. Deland, S.J. Grewe, S.M. Mastorides, Google AutoML versus Apple CreateML for histopathologic cancer diagnosis; which algorithms are better? 2019, arXiv preprint arXiv:1903.08057. http://dx.doi.org/10.48550/arXiv.1903.08057.

A. Choudhary, A. Choudhary, S. Suman, NLP applications for big data analytics within healthcare, in: S. Mishra, H. Tripathy, P. Mallick, K. Shaalan (Eds.), Augmented Intelligence in Healthcare: A Pragmatic and Integrated Analysis, Springer, Singapore, 2022, pp. 237–257, http://dx.doi.org/10.1007/978-98119-1076-0_13.

B. Ulrich, C. Grady, G. Demiris, T. Richmond, The competing Hum. Res. 4 (1) (2021) 25–31, http://dx.doi.org/10.1002/eahr.500076.

[16] N. Afzal, V.P. Mallipeddi, S. Sohn, H. Liu, R. Chaudhry, C.G. Scott, I.J. Kullo, A.M. Arruda-Olson, Natural language processing of clinical notes for identification of critical limb ischemia, Int. J. Med. Inform.111 (2018) 83-89http://dx.doi.org/10.1016/j.ijmedinf.2017.12.024.

[17] B. Galatzan, J. Carrington, S. Gephart, Testing the use of natural language processing software and content analysis to analyze nursing hand-off text data, Comput. Inform. Nurs. 39 (8) (2021) 411–417, http://dx.doi.org/10.1097/CIN. 0000000000000732.

[18] T. Tyagi, NeuraHealthNLP: An automated screening pipeline to detect undiagnosed cognitive impairment in electronic health records with deep learning and natural language processing, 2022, arXiv preprint arXiv:2202.00478. http: //dx.doi.org/10.48550/arXiv.2202.00478.

[19] M. Chowdhury, E.G. Cervantes, W.-Y. Chan, D.P. Seitz, Use of machine learning and artificial intelligence methods in geriatric mental health research involving electronic health record or administrative claims data: A systematic review, Front. Psychiatry 12 (2021) http://dx.doi.org/10.3389/fpsyt.2021.738466.

[20] Y. Juhn, H. Liu, Artificial intelligence approaches using natural language processing to advance EHR-based clinical research, J. Allergy Clin. Immunol. 145 (2) (2020) 463–469.

[21] T. Ahmed, M.M.A. Aziz, N. Mohammed, De-identification of electronic health record using neural network, Sci. Rep. 10 (1) (2020) 1–11, http://dx.doi.org/ 10.1038/s41598-020-75544-1.

[22] S. Wu, K. Roberts, S. Datta, J. Du, Z. Ji, Y. Si, S. Soni, Q. Wang, Q. Wei, Y. Xiang, H. Xu, Deep learning in clinical natural language processing: a methodical review, J. Am. Med. Inform. Assoc. 27 (3) (2020) 457–470, http: //dx.doi.org/10.1093/jamia/ocz200.

[23] H. Alzoubi, R. Alzubi, N. Ramzan, D. West, T. Al-Hadhrami, M. Alazab, A review of automatic phenotyping approaches using electronic health records, Electronics 8 (11) (2019) 1235, http://dx.doi.org/10.3390/electronics8111235.

[24] Y. Juhn, H. Liu, Natural language processing to advance EHR-based clinical research in Allergy, Asthma, and Immunology, J. Allergy Clin. Immunol. 145 (2019) http://dx.doi.org/10.1016/j.jaci.2019.12.897.

[25] T.A. Koleck, C. Dreisbach, P.E. Bourne, S. Bakken, Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review, J. Am. Med. Inform. Assoc. 26 (4) (2019) 364–379, http://dx.doi.org/10.1093/jamia/ocy173.

[26] Y. Wang, L. Wang, M. Rastegar-Mojarad, S. Moon, F. Shen, N. Afzal, S. Liu, Y. Zeng, S. Mehrabi, S. Sohn, et al., Clinical information extraction applications: a literature review, J. Biomed. Inform. 77 (2018) 34–49, http://dx.doi.org/10.1016/j.jbi.2017.11.011.

[27] Y. Luo, W.K. Thompson, T.M. Herr, Z. Zeng, M.A. Berendsen, S.R. Jonnalagadda, M.B. Carson, J. Starren, Natural language processing for EHR-based pharmacovigilance: a structured review, Drug Saf. 40 (11) (2017)1075–1089, http://dx.doi.org/10.1007/s40264-017-0558-6.

[28] A.K. Jabali, A. Waris, D.I. Khan, S. Ahmed, R.J. Hourani, Electronic health records: Three decades of bibliometric research productivity analysis and some insights, Inform. Med. Unlocked (2022) 100872, http://dx.doi.org/10.1016/j. imu.2022.100872.

[29] K. Ayre, H.G. Gordon, R. Dutta, J. Hodsoll, L.M. Howard, The prevalence and correlates of self-harm in the perinatal period: a systematic review, J. Clin. Psychiatry 81 (1) (2019) 15343, http://dx.doi.org/10.4088/JCP.19r12773.

A. Bittar, S. Velupillai, A. Roberts, R. Dutta, Text classification to inform suicide risk assessment in electronic health records, in: MedInfo, 2019, pp. 40–44, http://dx.doi.org/10.3233/SHTI190179.

[30] J. Downs, S. Velupillai, G. George, R. Holden, M. Kikoler, H. Dean, A. Fernandes, R. Dutta, Detection of suicidality in adolescents with autism spectrum disorders: developing a natural language processing approach for use in electronic health records, in: AMIA Annual Symposium Proceedings, Vol. 2017, American Medical Informatics Association, 2017, p. 641, URL https:

//www.ncbi.nlm.nih.gov/pmc/articles/PMC5977628/.

[31] H.D. Anderson, W.D. Pace, E. Brandt, R.D. Nielsen, R.R. Allen, A.M. Libby, D.R. West, R.J. Valuck, Monitoring suicidal patients in primary care using electronic health records, J. Am. Board Fam. Med. 28 (1) (2015) 65–71, URL https://www.jabfm.org/content/28/1/65.

[32] K. Ayre, A. Bittar, J. Kam, S. Verma, L.M. Howard, R. Dutta, Developing a natural language processing tool to identify perinatal self-harm in electronic healthcare records, PLoS One 16 (8) (2021) e0253809, http://dx.doi.org/10. 1371/journal.pone.0253809.

[33] B.E. Belsher, D.J. Smolenski, L.D. Pruitt, N.E. Bush, E.H. Beech, D.E. Workman, R.L. Morgan, D.P. Evatt, J. Tucker, N.A. Skopp, Prediction models for suicide attempts and deaths: a systematic review and simulation, JAMA Psychiatry 76 (6) (2019) 642–651, http://dx.doi.org/10.1001/jamapsychiatry.2019.0174.

[34] G.E. Simon, E. Johnson, J.M. Lawrence, R.C. Rossom, B. Ahmedani,

[35] F.L. Lynch, A. Beck, B. Waitzfelder, R. Ziebell, R.B. Penfold, et al., Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records, Am. J. Psychiatry 175 (10) (2018) 951–960, http://dx.doi.org/10. 1176/appi.ajp.2018.17101167.

[36] C.G. Walsh, J.D. Ribeiro, J.C. Franklin, Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning, J. Child Psychol. Psychiatry 59 (12) (2018) 1261–1270, http://dx.doi.org/10.1111/jcpp.12916.

[37] F.R. Tsui, L. Shi, V. Ruiz, N.D. Ryan, C. Biernesser, S. Iyengar, C.G. Walsh, D.A. Brent, Natural language processing and machine learning of electronic health records for prediction of first-time suicide attempts, JAMIA Open 4 (1) (2021) ooab011, http://dx.doi.org/10.1093/jamiaopen/ooab011.

[38] N.J. Carson, B. Mullin, M.J. Sanchez, F. Lu, K. Yang, M. Menezes, B.L. Cook, Identification of suicidal behavior among psychiatrically hospitalized adolescents using natural language processing and machine learning of electronic health records, PLoS One 14 (2) (2019) e0211116, http://dx.doi.org/10.1371/ journal.pone.0211116.

[39] G.K. Savova, J.J. Masanz, P.V. Ogren, J. Zheng, S. Sohn, K.C. Kipper-Schuler, C.G. Chute, Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications, J. Am. Med.Inform. Assoc. 17 (5) (2010) 507–513, http://dx.doi.org/10.1136/jamia.2009. 001560.

[40] D.J. Feller, J. Zucker, M.T. Yin, P. Gordon, N. Elhadad, Using clinical notes and natural language processing for automated HIV risk assessment, J. Acquir. Immune Defic. Syndr. (1999) 77 (2) (2018) 160, URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5762388/.

[41] S. Fu, G.S. Lopes, S.R. Pagali, B. Thorsteinsdottir, N.K. LeBrasseur, A. Wen, H. Liu, W.A. Rocca, J.E. Olson, J. St. Sauver, et al., Ascertainment of delirium status using natural language processing from electronic health records, J. Gerontol. Ser. A 77 (3) (2022) 524–530, http://dx.doi.org/10.1093/gerona/ glaa275.

[42] Y. Deng, J.A. Pacheco, A. Chung, C. Mao, J.C. Smith, J. Zhao, W.-Q. Wei, A. Barnado, C. Weng, C. Liu, et al., Natural language processing to identify lupus nephritis phenotype in electronic health records, 2021, arXiv preprint arXiv:2112.10821. http://dx.doi.org/10.48550/arXiv.2112.10821.

[43] M. Li, Z. Fei, M. Zeng, F.-X. Wu, Y. Li, Y. Pan, J. Wang, Automated ICD-9 coding via a deep learning approach, IEEE/ACM Trans. Comput. Biol. Bioinform. 16 (4) (2018) 1193–1202, http://dx.doi.org/10.1109/TCBB.2018.2817488.

[44] A. Kormilitzin, N. Vaci, Q. Liu, A. Nevado-Holgado, Med7: a transferable clinical natural language processing model for electronic health records, Artif. Intell. Med. 118 (2021) 102086, http://dx.doi.org/10.1016/j.artmed.2021.102086.

[45] S. Bird, E. Klein, E. Loper, Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit, O'Reilly Media, Inc, 2009, URL https://www.amazon.com/Natural-Language-Processing-PythonAnalyzing/dp/0596516495.

[46] C.D. Manning, M. Surdeanu, J. Bauer, J.R. Finkel, S. Bethard, D. McClosky,The Stanford CoreNLP natural language processing toolkit, in: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 2014, pp. 55–60, URL https://aclanthology.org/P14-5010.pdf. [49] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, et al., Transformers: State-of-the-art natural language processing, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, 2020, pp. 38–45, http://dx.doi.org/10.18653/v1/2020.emnlp-demos.6