

NLP/Deep Learning Techniques in Healthcare for Decision Making

Kamal Jain^{1*}, Vishal Prajapati²

¹Data Science Evangelist, Mentor, United Arab Emirates

²Department of Information Technology, G H Patel College of Engineering & Technology, Anand, Gujarat, India

Corresponding Author*

Kamal Jain

Data Science Evangelist, Mentor, United Arab Emirates

E-mail: kamaljain777@gmail.com

Copyright: 2021 Jain K, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Received 09 March, 2021; **Accepted** 23 March, 2021; **Published** 30 March, 2021

Abstract

Day by day the health care domain is generating millions of records of patients in a structured and unstructured way. By applying deep learning techniques, it can be converted into a well-structured form i.e. Electronic Health Record (EHR). Decision making is one of the key parts in the Healthcare domain. In the decision making process doctors must refer to many data like laboratory reports, diagnosis reports, medical images, demographic information about the patient, clinical notes, and map it collectively with the concepts of medical science. Here AI especially Natural Language Processing and deep learning can be helpful in many ways.

The objective of this article is to depict the importance of Deep Learning and Natural Language processing EHRs. Based on EHR, the doctor can take quick decisions in case of an emergency. Apart from that, it can also be effective in clinical predictions, detect disease at an earlier stage, forecasting future need of regular check-ups, predictions of hospitalization soon if required. The provision of an enormous amount of clinical information particularly EHR has stimulated the expansion of deep learning techniques that assist within the rapid analysis of patient data.

Keywords: Natural language processing, NLP, Deep learning, Data science, Artificial intelligence, EHR, Healthcare, Electronic health records

Introduction

The ubiquitous adoption of electronic health records in hospitals and other healthcare facilities generates vast real-world information, which is very valuable for conducting clinical research. Over the past many years, electronic health records (EHR) systems are being widely adopted across clinics and hospitals. Analysis of this huge data is the foundation to provide improved healthcare to the patients. However, manual review of this vast amount of data generated from multiple sources is costly and very time-consuming. It brings huge challenges while attempting to review this data meaningfully. Hence, the role of artificial intelligence (AI) techniques is becoming important in enhancing clinical research and care. As a large number of EHR are locked in clinical narratives, Natural Language Processing (NLP) and Deep Learning (DL) techniques have been leveraged to extract information. While computer vision techniques can be used for medical imaging, NLP can be used for analysing unstructured information in the EHR, reinforcement learning techniques can be used in the context of robotics-assisted operations. NLP algorithms can be used to identify clinically relevant phenotypes while analysing text and determining the grammatical relationships between phrases. Rule-based NLP techniques can be used to get high sensitivity (identification of a large proportion of true cases) and high positive predictive value in clinical records.

Healthcare is one of the domains where computer science is becoming very supportive of varied tasks. AI is increasingly being adopted across the healthcare industry ranging from basic level practices to specialization, and lots of the foremost exciting AI applications leverage language processing (NLP). The capabilities of these AI techniques potentially enable the identification of distinctive clinical characteristics among patients which further helps in clinical care and minimizing methodological heterogeneity in medical research concerning various health diseases.

What is Natural Language Processing?

Natural language processing (NLP) focuses on analysing text and speech to infer meaning from words. Recurrent neural networks (RNNs) - deep learning algorithms play a key role in processing sequential inputs like language, speech, and time-series data [1]. Deep learning is a subset of machine learning having the capacity of learning unsupervised data from unstructured or unlabeled data. On the other side deep learning will be able to learn optimal features from available data without human intervention.

Natural language processing is used to describe the process of using computer algorithms to identify key elements in everyday language and extract meaning from unstructured spoken or written input. NLP requires skills in artificial intelligence, computational linguistics, and other machine learning disciplines [2].

There are two approaches to NLP.

1. A rule-based approach where computer follows pre-defined rules in the program
2. Machine Learning-based approach where we can have supervised and unsupervised learning methods. In supervised learning, the computer learns latent rules through human guidance called annotation while in unsupervised learning there is no human interaction.

NLP algorithms, first extract information or concepts from EHRs, then process extracted information, and finally classify patients into a subgroup as per rules and learners. Procedures for NLP are complex because it consists of multiple techniques together. NLP will map phrases or words to concepts of interest, and it needs careful pre-processing of text and it will be converted to document from the natural form [3].

A few examples of low-level NLP tasks (text pre-processing) include the following:

1. Sentence boundary detection, which is usually defined by a period
2. Tokenization (breaking a sentence into individual tokens)
3. Stemming (reducing word into a root form)
4. Lemmatization (process mapping a token)

A few examples of higher-level NLP tasks include the following:

1. Named entity recognition
2. Setting up negation rules

The Figure 1 depicts the text pre-processing and classification of an asthma patient.

Role of Deep Learning/NLP in Healthcare

Basically, in terms of users, we can categorize the healthcare information coming from below four sources -

1. Doctors
2. Patient
3. Paramedical staff
4. Pharmaceuticals

Every post-process depends on the diagnosis of a disease. If the disease is identified properly then one can get proper treatment. Sometimes the situation of patients gets serious due to delays in decision making (Table 1).

Figure 2 represents, deep learning is an approach where there is a large scale network that is going to accept a variety of input data types like text, image, audio, time-series data, etc. for each data type learns a useful features in its lower level towers. The data from each pillar is then merged and flows through higher levels, allowing the Deep Neural Network to reach to conclusion based on reasoning and evidence across data types.

Natural language processing can help healthcare in information extraction, unstructured data to structured data conversion, Document categorization,

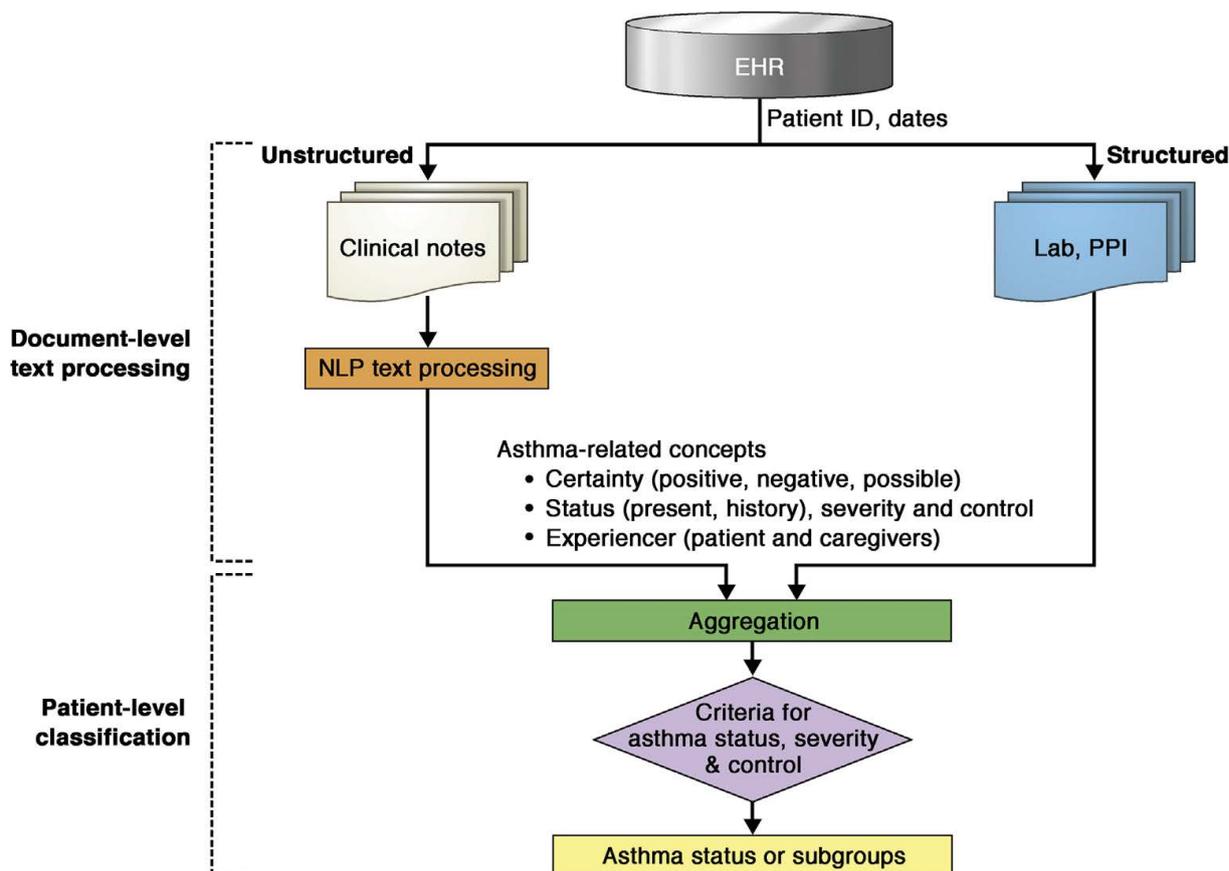


Figure 1: NLP Algorithms for document-level text processing and patient-level classification [3].

Table 1. Disease stage, process required and the actions to be taken.

Sr. No.	Situation	Process	Tools/Devices/Reports/Actions
1	Pre-Disease	Health Awareness	Smart sensor devices to provide health updates
		Regular health check-ups	Devices for health check-ups
		How to stay away from disease?	Training programs and devices for various exercises
2	In Disease	Diagnosis	Blood sample, X-Ray Image, Urine Sample, Biopsy, Devices for checking vital signs in the body (Oxygen, sugar, heartbeat, pulses, etc.)
		Treatment	Pre-surgery treatments, Minor Surgery, Major surgery, Combined surgery, Post-surgery treatments
		Medication	Drug dosage based on the nature of the body for a single disease, multiple diseases
3	Post Disease	Follow-ups	Prediction of the number of follow-ups or time period of follow-ups required
		Improvement comparison	Compare Pre, In and post disease reports
		Future predictions	

and summarization. Ultimately it will reduce administrative cost by efficient billing and accurate prior authorization approval. It also going to create Medical Value by providing support for ineffective clinical decisions and streamlined medical policy assessment etc. [4,5]. Apart from that it will help in improving patient interactions with the provider and the EHR, increase patient health awareness, improve care quality, and identify patients with critical care needed.

In healthcare, sequential deep learning and Natural language technologies boosts applications within domains like electronic health records (EHRs) [1,6].

An outsized medical organization can capture 10 million patients' medical transactions in a decade. One hospital alone generates 150,000 pieces of data. Aggregation of all EHRs of this scale exemplifies 100 million years of patient data and 200,000 years of doctors' practice, and it covers excess of rare conditions and diseases [1].

The Figure 3 depicts various stages of processing

- Unstructured EHR data. Medical records are stored in heterogeneous formats as per the working mechanism and data structures of a specific hospital. So, it will vary from one hospital to another.
- Data standardization. Based on FHIR, data will be standardized into the same or we can say homogeneous format by mapping data from multiple sources

- Sequencing. By temporally sequencing all data into a patient timeline, time-based deep-learning techniques can be applied to the entirety of EHR datasets for making predictions about single patients.

NLP can be helpful in the neurology domain too. The Figure 4 depicts how the iterative process of the NLP algorithm which can be helpful in identifying 24 brain imaging phenotypes in two areas of NHS Scotland which had excellent positive predictive value for cerebrovascular and neurodegenerative phenotypes.

The NLP system, Edinburgh Information Extraction for Radiology reports (EdIE-R), is a staged pipeline process as depicted in the Figure 4, with XML rule-based text mining software at its core.

Conclusion

NLP will play a significant role in accelerating the decision-making process in the healthcare area. However, the real rewards of developing good algorithms will depend heavily on the quality of the data that they acquire and maintain. The faster decision-making process will allow physicians to focus on the value-added care of patients. NLP with Deep Learning and Computer Vision can process a variety of data together to take precise decisions. Collaborative research can lead to a higher level of treatment in Healthcare. Considering the impact of AI techniques, these systems need to be designed and built very carefully in a larger socio-ecological context of clinical care settings to provide better healthcare to society.

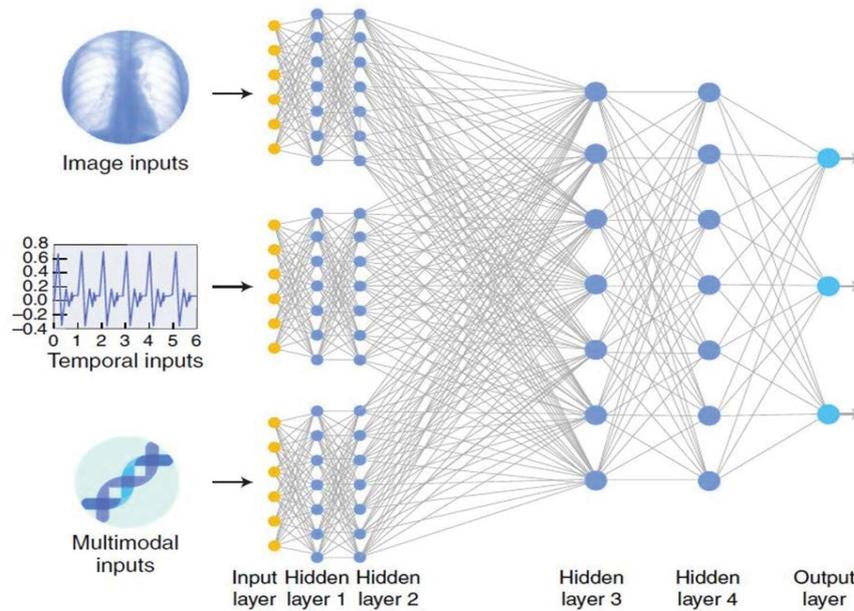


Figure 2: Deep learning can detect features and learn from a variety of data types [1].

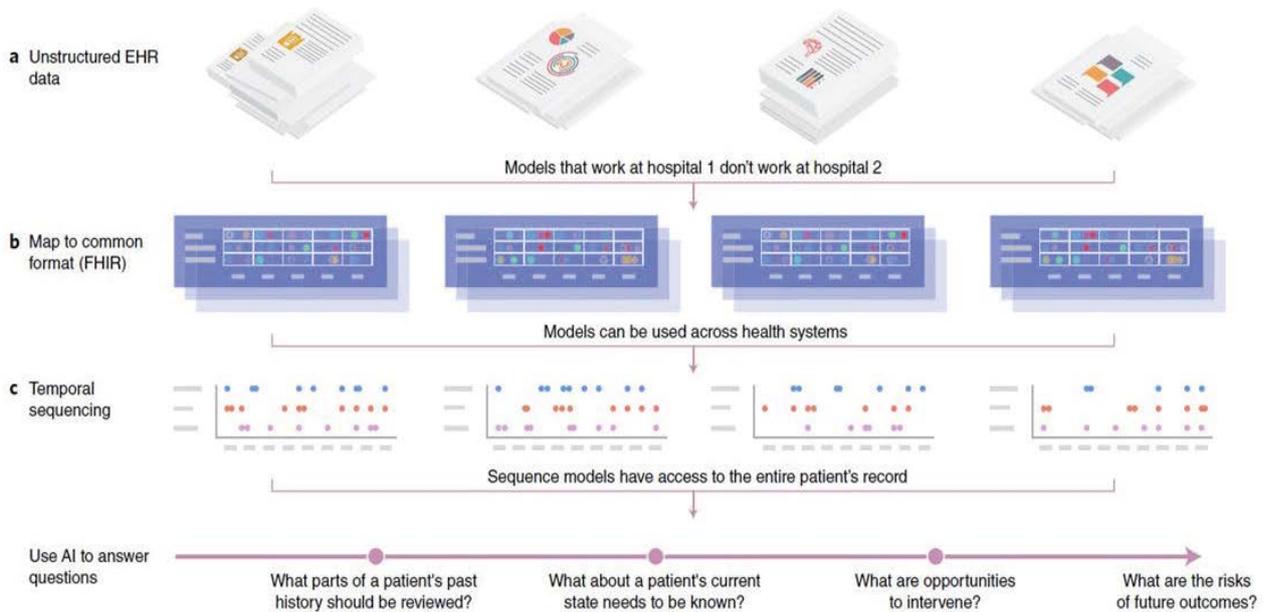


Figure 3: Making predictions using EHRs [1].

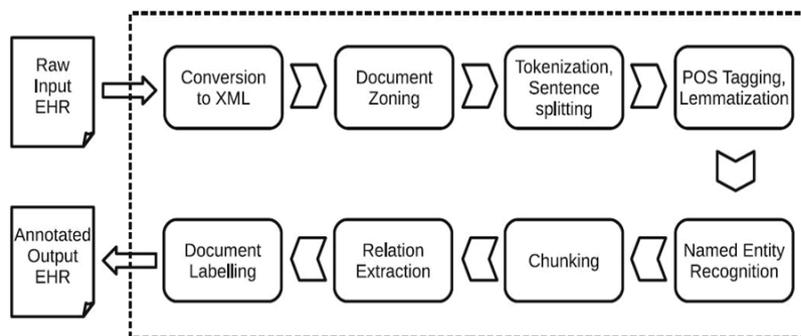


Figure 4: EdIE-R System Architecture [7].

References

1. Esteva, A., et al. "A guide to deep learning in healthcare." *Nature Med* 25 (2019):24-29.
2. Bresnick, J. "What Is the Role of Natural Language Processing in Healthcare?" *Health IT Analytics* (2020) Retrieved from <https://healthitanalytics.com/features/what-is-the-role-of-natural-language-processing-in-healthcare>
3. Juhn, Y., & Liu, H. "Artificial intelligence approaches using natural language processing to advance EHR-based clinical research." *J Allerg Clin Immunol* 145.2 (2020):463-469.
4. Rangasamy, S. et al. "Natural language processing in healthcare." *McKinsey & Company* (2018) Retrieved from <https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights>

5. "Top 12 Use Cases of Natural Language Processing in Healthcare" Retrieved from <https://marutitech.com/use-cases-of-natural-language-processing-in-healthcare/>
6. Poongodi, T., et al. "Deep Learning Techniques for Electronic Health Record (EHR) Analysis." In: Bhoi, A., et al. (eds) Bio-inspired Neurocomputing. Studies in Computational Intelligence, vol 903. Springer, Singapore.
7. Wheater, E., et al. "A validated natural language processing algorithm for brain imaging phenotypes from radiology reports in UK electronic health records." *BMC Med Informatic Decision Making* 19 (2019):184.