International Journal of Computer Applications (0975 – 8887)
Volume 174 – No. 24, March 2021

Progress in Deep Learning Mechanisms for Information Extraction from Social Networks: An Expository Overview

Israel Fianyi
University of Tasmania
School of Information and Communication Technology
2 Invermay Rd, Launceston, Tasmania, 7248

Gifty Andoh Appiah
Westminster University, London,
Unite Kingdom

ABSTRACT

Deep learning algorithms have shown to be robust in extracting high quality information from a wide range of online platforms. Incidentally, social networks and other related online platforms are known to hold a copious amount of unstructured user-generated content. To date, machine learning and deep learning approaches for mining textual data have received so much attention from researchers and industry players Deep learning is good at independently learning from complex feature representation and make intelligent decisions from data. However, with the influx if different deep learning methods for information extraction, understanding and finding the current challenges and recent advances in these algorithms is daunting. This paper investigates existing pieces of literature to appreciate the trajectory of deep learning for information extraction in Natural Language Understanding. The study further considers the state-of-the-art, open challenges, as well as the tools and methodologies involved in undertaking information extraction tasks from Unstructured data. The study considers relevant published articles from the year 2009-2020 that focused on deep learning approach for information extraction from text. The investigations of this paper provide extensive clarity to the research field of Natural Language Processing with deep learning. It identifies current research problems, recommends directions for future research. The paper is designed to help non-expert researchers comprehend the fundamentals of deep learning and Natural Language Processing methods for Information Extraction.

General Terms

Machine Learning, Deep learning, Pattern Recognition, Algorithms, Natural Language Processing, Information Extraction, Social Networks

Keywords

Deep learning, Algorithms, Natural Language Processing, Information Extraction, Social Networks

1. INTRODUCTION

The association between data and social events is strengthened by the positive results that are attained when data is analysed and used in making informed decisions. Humans are noted to communicate with audios, text, and can ascertain the meaning of an assertion by evaluating the context in which they are communicated [1, 2]. Consequently, with the advent of the Internet, particularly social media, textual means of communication became extremely rampant. Incidentally, these textual corpora has become some of the vital data types, including others such as recorded voices or speech, images, videos, audios and symbols [3, 4]. Textual data are peculiar to the element or subject area they are used for, such as a domain, language and among others. This is due to the dynamic nature of the human language which has subsequently contributed to the massive iterative textual and video interaction that goes on over social media today, parenthetically, there is a plethora of information on the social media platforms [5, 6]. The accessibility of these enormous data accumulated from various sources has facilitated the surge of interest in approaches to make value and extract high-quality relevant information from these copious data.

Although Natural Language Processing and its related technologies have made it possible for the information to be extracted from online platforms, there have been various challenges associated with the techniques used to extract relevant information. This study surveys the progress in open information extraction techniques from social networks, blogs and textual news media with their peculiar applications. The adapted methodology for this paper is discussed in the next section.

2. METHODOLOGY

Drawing inspiration from Beel, Langer [7], [8, 9], this paper searched various journal databases (Scopus, Web of Science, ScienceDirect, IEEE Xplore) with specific keywords and their synonyms. The method of conducting surveys has been successfully used in various studies in different disciplines [10, 11]. **Keywords:** (**"machine learning"** OR **"deep learning"** OR **"artificial intelligence"**) AND (**"information extraction"** OR **"information retrieval"**) AND (**"social networks"** OR **"social media"**))

Additionally, the following questions underpinned the motivation for this paper:

1. What are the current machine learning approaches for extracting information from social networks?
2. How does the deep learning classifiers work for information extraction task?
3. What are text classification task categorised as information extraction problem?

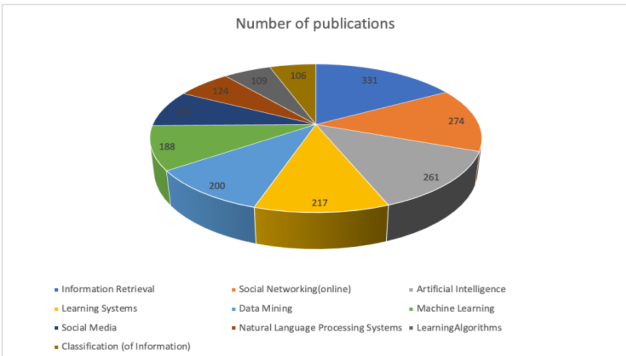

Figure 1.The categorical outcome of keyword search from databases

Figure 1 shows the segmental outcome of the keyword search from the various database, with 331 publications on data mining, followed by 274 publication on social networks representing the highest of publications retrieved from the search. The least publication with 106 documents belonged to the classification of information publications. Furthermore, Figure 2 below represents the distribution of years the various documents were published. From 1990 through to about 2003, there were no documents published within the purview of the query terms. However, there was a rise in the publication from 2004, then saw about 120 publications in 2019 then torpedoed in the year 2020 with about 40 publications, though the year 2020 is yet to come to an end.

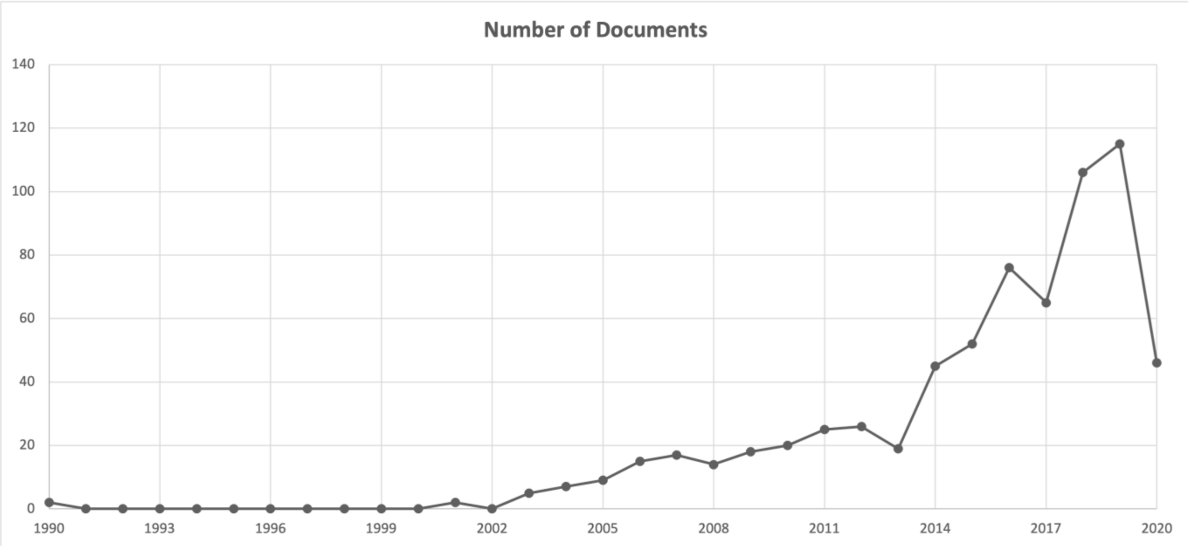

Figure 2. Distribution of the Years of Publications

3. RELATED WORKS

3.1 Natural Language Processing

The recent explosion in Social Networks has facilitated the massive research underway with regards to Natural Language Processing (NLP). This is because there is a plethora of information that is recorded on social media by humans in the form of text, video and audio. Sheth [12] in a study estimated that over 25billlion of data are generated daily on social media. The copious nature of the data inflow makes it humanly impossible for anyone looking for information from these data to undertake this task manually. Natural Language Processing, with its related technology, has made it possible for the information to be extracted from these massive data and subsequently make value out of them. Xu and Tan [13], indicated in their study that, data available on the web is usually ambiguous and unstructured. As such, many research and studies have been undertaken to transform these unstructured data into structured. This is attained by manipulating the data structures and redefining the rules that need to be followed in order to make the data structured. Subsequently, Maynard, Bontcheva [14] argues that one of the disadvantages associated with data structuring is the fact that some valuable information is truncated due to the predefined rules to get the data structured. Lim, Tucker [15] substantiated in their study of using unstructured data to achieve a specific objective of their research. They mentioned, using structured or unstructured data sources have both their merits and demerits. However, one would have varied options of data in an unstructured data source to make value from it. Recently, NLP has been successfully applied in areas of information extraction, question answering, intelligent systems and among others [16-18]. The social media provides the platform for massive information to be recorded, stored, shared and many other related activities [19, 20]. Consequently, there are mammoth data available on these varied social media platforms to be used for the solutions to several problems of society given the right tools for data mining such as machine learning and its related technologies. *In the subsequent sections, the paper discusses various social network platforms and their related application.*

3.2 Unstructured Social Data Sources

Textual communication has become a rampant medium of human communication over social media, as suggested in Tarasconi, Farina [21]. They further defined unstructured data as exiting data that are unlabelled or unorganised and have no predefined models. Salloum, Al-Emran [2], [22] also asserted, textual data is a generally accessible source of data, particularly on the social networks including other data types such as images, video, audios and among others. Musial, Brodka [23] categorised textual data into style, domain and languages. Incidentally, terminologies used in textual data extraction are domain oriented. With regards to social media and other related web platforms, several domains are classified as unstructured, that can be collected and analysed to attain a specific goal in a given domain. The next section discusses various social data sources.

3.3 Patents

Kang, Souili [24] defined patent as the arrangements between an inventor and a government regarding granting the innovator the legal rights to use, reproduce and trade the invention within a given period of time. These documents, known as a patent, give the necessary details of how the invention works and its associated rudiments. Thomas and Sangeetha [25] asserted that a patent is considered an unstructured data source because they are raw and consist of much-hidden information. Despite the availability of several patent collection and search applications. Information can be extracted in these patents, regarding the purpose, the general importance of them as well as the possible future usage [26]. Additionally, patents documents are constituted by various subjects, making it a multi-domain document for proper and respective information extraction. They also contain figures and tables which are usually excluded from textual information processing [24, 27].

3.4 Publications

The advent of automatic text processing and information extraction was implemented and used in identifying published articles [28]. Publications encompass journals, conference documents, books as wells other related academics papers as posited in Sato and Sato [29]. Information extraction applications became extremely relevant in publication materials due to the multiplicity of technical, voluminous and systematics publications that are released annually. Subsequently, they are easily attributed to a domain without a challenge [30]. The structure of publication has made it possible to identify solutions for problems in their respective domains. The abstract, index terms constitute a typical publication, introduction through to methods, results and most original studies have future research and recommendation. The imperativeness of extracting information from publications lies in their resources to provide verifiable and repeatable methods of a study, which is usually evidenced in their results. For instance, Ali, Hussain [31] mentioned the possibility of extracting medical knowledge from publications, particularly of drug discoveries and its associated scientific approaches. Additionally, Tourassi, Yoon [32] mentioned the advantages of web mining offers, especially in building knowledge ontology in epidemiological outcomes.

3.5 Blogs

Blogs are online-based discussion or information storage platform similar to conventional diaries and notice boards. Jebbor and Benhlima [33] opined, the recent explosion in blogs require that requisite rule-based mechanisms are employed to extract relevant information from these blogs. The authors further posited, most blogs are professionally written and contain vital intelligence that can be tapped. Consequently, Saad and Mansor [34], using Named Entity Recognition in blogs, identified and developed an archetype system model in crime news in the Malay language. As a result, they obtained nearly 80% of precision when benchmarked and applied in real-world scenarios. The ability of users to express their opinions in a discussion via comments on blogs makes it a robust data source for intelligence building. Klein, Altuntas [35], in a knowledge-oriented approach, extracted the sentiments of investors from blogs. The success of the extraction was seen in investors attesting to the algorithm accurately predicting their views within the given time frame. Additionally, in exploring the dexterity of blogs in facilitating ample information storage, a recent study by Abrantes and Cordeiro [36] in using web-based blogs to extract information on the adverse effects of drugs from the comments of users, who have lived and or living on drugs. There is a plethora of research in harnessing the potential of blogs regarding information storage, extraction and intelligence building.

3.6 Social Media

The Social media encapsulate various online social networks such as Facebook, Twitter, YouTube, Youku, Google+ and the like [37, 38]. These platforms are characterised by users being able to publish their personal information, news, sentiments as well as their thought. It is accessible by the groups on the platform or the public, if users choose to make it public [39, 40]. The application of Natural Language Processing and machine learning paradigms have thrown more light on the propensity of using these social media platforms to promote human-centered computing and its related technologies. Although text mining in social media has been around for a while, Beigi, Hu [41] used social media outlets to unravel the sentiments expressed in disaster relief situations. Additionally, Pandey and Natarajan [42] elaborated on the advantages the social media offers concerning the occurrence of disasters, mainly because they occur mostly unannounced. Speaking to the ubiquitous nature of social media makes the spread of information regarding sending signs of danger and possible untoward to the public cost-effective and relatively faster. In a more technical application, Luong, Cao [43] developed an algorithm using social media text to predict the intentions of customers regarding patronage of goods and services as well as feedback after the purchase. Furthermore, Ramírez-Cifuentes, Mayans [44] used social media data to determine the risk associated with the early detection of anorexia. Likewise, Nikfarjam, Sarker [45] established the success of pharmacovigilance from mining social media data. The study revealed the ability of social content to assist in determining the adverse of manufactured drugs in users. Although social media seems to be making immense progress in Artificial Intelligence, there remain enormous challenges in using social media data. Unlike blogs and publications, the multi-domain nature of social media makes it daunting to determine the requisite and relevant text for a domain in view [46-48]. The proliferation of emojis, jargons, abbreviations and short handwriting in social media text contribute to the complexity and challenges in mining social media text relative to the others [49-52].

4. SOCIAL NETWORKS DATA COLLETION

The quality of social media textual mining is greatly incumbent on the processes that are undertaken to retrieve these data as suggested in Fabian, Baumann [55]. The quality of the input data can significantly determine the results of an analysis from any of the social network outlet. Since the upsurge in text mining with Natural Language Processing ensued, various challenges have been identified to impede the process and quality of data collection from social networks [56-58]. Chen, Li [59] established, one of the significant challenges online social network faces in recent times has been an overload of information as well as information pollution all together known as information rumours (misinformation). The social networks have become an integral part of people's daily activity, making it a good source of data. Nevertheless, most of these online social network's contents are constituted by excessive rumours, and the appropriate data collection methodology is essential in extracting accurate and relevant information from social networks [60]. Y. F. Chen, Li, Liang, and Qi (2018) developed a methodology for data collection and annotation that defuses rumours in the Social Networks, to extract quality relevant information eventually. Abid, Ameur [61], proposed a methodology for data collection and annotation that identifies distrustful content and malicious contents. The study was successful in collecting and identifying suspicious contents from various social networks. Information extraction applications are domain directed; as such, they determine the type of textual data that is collected from the intended online and social network platform. Zaghouani and Charfi [62], on the other hand, proposed a framework to defuse the multi-lingual challenge that confronts NLP applications. It is a multi-corpus data collection and annotation paradigm to undertake information extraction. Multiple data sources have proven to facilitate the robustness of any information extraction applications in Natural Language processing [63, 64]. *In the next section, we discuss recent trends (deep learning methods) for information extraction.*

5. DEEP LEARNING OVERVIEW

Deep learning embodies multilayered features that enable it to learn from data representation hierarchically [65, 66]. Deep learning facilitates spontaneous feature wrangling without human participation and is inherently robust at learning complex feature representation, as asserted in Zhou, Zeng [67], [68]. A traditional machine learning assumes a model which can be best described as Shallow Learning (SL), Including Naive Bayes, Decision trees and Support Vector Machines (SVM) as suggested by Liu [69]. A research paper by Mao, Feng [70], [71] implied, these conventional approaches are single layered. Additionally, Marcus [72], [73] opined, to attain significant and relative success in any deep learning application, it requires deep architecture. This deep architecture is composed of multiple layers; consequently, machine learning methods that use at most two layers are considered as shallow learning representation. Zhang, Yao [65], [74, 75], mentioned, due to the adaptability nature of deep learning, the conceptualisation of the layered models becomes relatively easy. Subsequently, deep learning is defined as the technique that is embedded with feature space that processes an observation one after the other, and this referred to as neurons. Deep learning, in recent times, has proven to be a relatively exceptional tool for Natural Language Processing (NLP) and its related field, such as information extraction. A well-characterised text mining in a given feature space is usually in hidden layers, the association between a semantic and its relative context can be learned with this approach. In the following section, we give a breakdown of the various components of deep learning for text processing.

5.1 Neural Networks

Deep Neural Networks (DNN) is a notable example of deep learning, and the term is deduced from the nature of the reasoning characteristics of the human brains as defined by Zhao, Yang [76]. The human brains are constituted by a compactly interconnected collection of nerve cells. These cell nerves known as neurons are fundamental information processing unit in the human brains, and there are over hundred (100) billions of such neurons and about hundred (100) trillions of connections between these elements [77]. Concurrently, using the neurons makes the human brains perform their functions faster, making biological neurons able to learn through experience. Qi, Das [78] indicated, the adoptions of biological neural networks into computing, known as Artificial Neural Networks(ANN), has in recent times contributed massively to the success of artificial intelligence and its related technologies. Figure 3 and Figure 4, typically shows the features of Biological Neurons (BN) and the Artificial Neural Network, respectively. Subsequently, Table 1 illustrates a comparison between ANN and biological neurons.

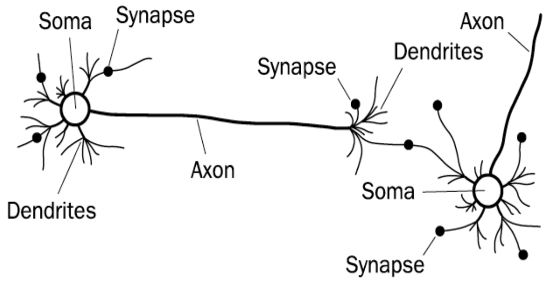

Figure 3. Exemplification of the Human Neural Networks

Every neural network has an input layer, which receives initial data for processing. The data then goes through a hidden layer where the processing and mechanisation take place. Subsequently, the output layer of the neural network is the final stage, where an outcome of the data inputted, processed is displayed based on the expected outcome and the characteristics of the data inputted. In the subsequent section, we discuss the various artificial neural network applications and their respective classifiers.

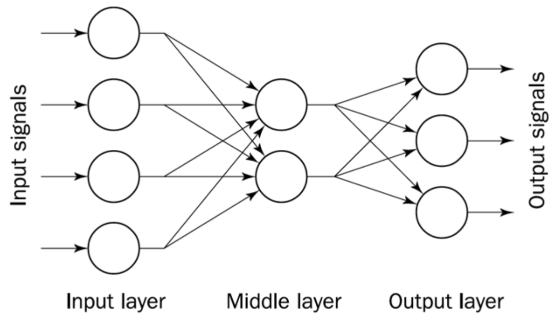

Figure 4. Typology of Artificial Neural Networks

Table 1.Comparing Artificial and Biological Neural Networks

Biological Neural Network	Artificial Neural Network
Soma	Neuron
Dendrite	Input
Axon	Output
Synapse	Weight

5.2 Applications of Artificial Neural Network in NLP

Identifying the appropriate artificial neural network tools and their classifiers for a respective task is imperative.

Table *2* below shows a categorisation of the three neural networks types with the corresponding task they can conveniently be used.

Table 2.Categorical analysis of ANNs and their corresponding applicable, suitable task

ANN	Data types	Applications	Reason for appropriateness
RNN	Textual, Audio data. Classification & Regression rediction problems as well as Generative models	Language modelling and prediction, speech recognition, general NLP	Ability to use both current input and also memorise the previous input- due to its internalised memory. e.g. **LSTM**
CNN	Image data, classification and Regression extrapolations	Image recognition and classification	Due to its high accuracy to use current inputs. Ability to generate a dimensional duo image with an Internalised representation
Multilayer Perceptron, MLPs	Image, textual & time-series data Issues on (image, textual and time-series data	Classification, regression prediction	Ability to acquire knowledge in inputs and outputs for mapping. Suitable for the tabulated dataset, e.g. (CSV or Excel files)
Hybrid Neural Networks	Audio, video, textual & time-series data	Clustering, classification, Regression, General NLP, image recognition	Integration of two or more neural networks

Source: Author(s) |RNN: Recurrent Neural Network| CNN: Convolutional Neural Network

5.3 Attention Mechanism and Memory Networks

Attention mechanism and memory networks stemmed from the cognition activity of humans, which is an intentional ability of humans to visually focus on a particular thing while ignoring others within the same environment [79, 80]. In the information extraction task, attention mechanism does not encode the entire sources of a text or sentence into a specified vector. Instead, it permits the decoder to "attend" to various aspects of the source of the text or sentence at every given phase of the output creation [81]. Nevertheless, in a CNN and RNN encoder-decoder framework, there is usually prospective irrelevant information generated by the encoder, mainly when the input information is rich. The solution to this problem has been to use the attention mechanism, which lets the model focus on an input item. Thus, the model chooses the part of the text or sentence to "attend" to [68]. Recently, Google developed a memory-based application known as Bidirectional Encoder Representation from Transformers (BERT); A language model that can predict the probability of the occurrence of a word, based on the examples of previous textual information. Implying that this application does not use the RNN framework encoders instead, the attention mechanism transformers [82].

6. RESULTS AND ANALYSIS

Table 3 below represents the outcome of keywords on various information extraction, techniques, methods used for the study. The essence of this table is to demonstrate to early career researchers and non-expert in the field to understand how to search for relevant materials for a study in information extraction. In any given research project, a chosen method is required as well as the dataset needed for the task. Table 3 shows various artificial neural network methods adopted for a project, types of datasets, number of sources employed, the study outcomes and the types of publication outlets used in reporting the results of the projects.

Table 3. Information Extraction Applications

Authors	Year of Publication	Methods used	Data source	Number of data sources	Study Outcome	Reference type
[83]	2013	Feature vector generator	Twitter datasets	1	Automatically named entity aliasing resolution (ANEAR)	Conference Proceedings
[84]	2019	Comprehensive framework	Journals	NA	Fake News: Fundamental theories, detection strategies	Conference Proceedings
[67]	2018	Deep sentiment hashing model	Twitter, Facebook, Instagram	3	Sentiment-based text retrieval	Journal Article
[85]	2019	the bi-directional neural	review datasets	2	event detection	Journal Article

		language model				
[68]	2019	Deep Recurrent Convolution Neural Networks	Twitter datasets	1	Subjectivity Classification	Journal Article
[86]	2018	Contrast Rule-based Sentiment Analysis algorithm	Amazon datasets	1	Multi-level Sentiment Information Extraction	Conference Proceedings
[87]	2018	Classification through NB (Multinomial naive Bayes)	Twitter stream	1	the categorisation of users into extremist and non-extremist	Conference Proceedings
[88]	2016	Topical Mesh representation	drug review corpora	15	Improve the performance of biomedical document retrieval and classification tasks	Journal Article
[89]	2015	Scalable classifier feature extraction	Twitter stream	1	Detect personal health mentions on Twitter	Journal Article
[90]	2013	logical sitemap mining	menus, breadcrumbs, sitemap	3	Building enhanced link context	Serial
[91]	2017	Deep learning techniques	Twitter stream	1	Chatbot for customer service on social media	Conference Proceedings
[92]	2016	Matching model	Ubuntu chat rooms (human-human dialogues)	1	Topic retrieval-based chatbots	Journal Article
[93]	2016	computational intelligence approaches	Twitter stream	1	The influence of irony and sarcasm on NLP techniques	Serial
[94]	2016	multiple instances learning approach	News article	1	Identifying key sentences and detecting events	Conference Proceedings
[95]	2016	Multi-representation approach	10-K reports (US trading companies)	1	Text regression of financial risks	Conference Proceedings
[32]	2016	Web crawling & tailored information extraction procedures	online obituary articles	1	Web mining for cancer-related epidemiological discoveries	Journal Article
[96]	2017	Long Short-Term Memory(LSTM) based models	The community question and answers corpus	1	Chinese lexical normalisation	Serial
[97]	2018	Conditional Random Field (CRF) based model & deep learning model	Electronic Health Records	1	Information extraction of Chinese electronic health records	Conference Proceedings
[98]	2017	a scalable distributed linguistic analysis pipeline	RSS feeds, and news articles	2	Real-time event recognition	Conference Proceedings
[21]	2018	Social Big Board	Twitter stream	1	real-time disaster-related social media monitoring	Conference Proceedings
[99]	2019	Partition protocol (K10)Ê	Twitter datasets (health)	1	feature enrichment on healthcare text classification system	Journal Article
[50]	2016	Standard natural language processing	geotagged Twitter text	1	Facility detection and popularity assessment	Conference Proceedings
[100]	2018	neural network models	Stock Twits tweets	1	Finding expert authors in the financial forum	Conference Proceedings
[44]	2018	domain-specific vocabulary model	Twitter steam (user’s writings)	1	Early risk detection of anorexia	Serial

6.1 Recurrent Neural Network Models for NLP

Table 4 shows various projects that adopted RNN classifiers for information extraction task. Recurrent neural networks (RNNs) are characterised by their ability to adapt when changes in data pattern occur over time. Subsequently, the same layer is used for the input in every given step, while using the output. It implies that the output of the previous becomes the input of the next step [85]. Recurrent Neural Networks have predominantly gained prominence in sentiment analysis related applications, However, in recent times, it is being explored and applied in other exciting domains. Comparing RNN to CNN models which are discussed in the subsequent section. Recurrent Neural Networks have more flexible computational and statistical phases. These attributes make the capturing of contextual language dependencies in various text length possible [86, 87]. The table is designed to be self-explanatory and informative for anyone that seeks to understand how RNN has been used over the years for information extraction tasks. Another notable element of the table below is the evaluation metrics used in estimating the robustness and accuracy of the application.

The experiments are characterised by Accuracy, Precision, Recall rate and the F1-measure to evaluate the performance of the text classification model. The following equations are used in calculating the indicators:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} X100\% \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} X100\% \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} X100\% \quad (3)$$

$$F_1 = \frac{2\ X\ precision\ X\ recall}{precision + recall} \quad (4)$$

Table 4.Recent Applications of RNN model in Information Extraction Projects

No.	Study& Author(s)	Domain Application	Data set	Language	Performance	Learning Type	Model
1	**Event -text Extraction**						
	Viani, Miller [104]	Medical	75 cardiology reports	Italian English	F1 score:90.1% Precision:88.6% Recall: 91.7%	Supervised	RNN+KBS
2	Gupta, Pawar [105]	Pharmacovigilance	Twitter	English	F1 score:74.7% Precision:72.7% Recall: 78.0%	Semi-supervised	RNN
3	Gao, Young [106]	Medical (cancer)	942 de-identified pathology reports	English	F1 score:85.2% Precision:76.8% Recall: NA	unsupervised	HAN+RNN
	Text Summarisation						
4	Xiang, Xu [107]	Textual Semantics	Daily Mail and DUC-2004 datasets	English	F1score: 42.0% Recall: 33.9%	Supervised	MLP+RNN+CNN
5	Liu, Li [108]	Social networks	web pages	English	F1 score:85.2% Precision: NA Recall: NA	unsupervised	HAN+RNN
	Question Answering						
6	Yin, Chang [109]	user intention identification	internal advertising dataset	English	F1 score:79.8% Precision: NA Recall: NA	Supervised	RNN
	Information extraction from Image						
7	Banerjee, Ling [110]	Medical	7370 tomography (CT)	English	F1Score :93% Precision:91% Recall:95% (UPMC dataset)	supervised	DPA-HNN +CNN
	Named Entity recognition						
8	Banik and Rahman [111]	Linguistic	Online Newspaper	Bangla	F1 score:69.0% Precision:93.3% Recall:99.7%	Supervised	RNN+GRU+Manual annotation
9	Zhang, Cui [112]	Relational	SemEval-	English	F1 score:82.9%	Supervised	Multi-

		Semantics	2010		Precision:84.0 Recall:82.0		gram CNN+RNN
10	Chowdhury, Dong [113]	Medical	Electronic Medical Record	Chinese	F1 score: 89.37 % Precision:89.31 % Recall: 89.47%	Supervised	Bi-directional RNN
11	Nagesh Bhattu, Satya Krishna [114]	Linguistic	Indian text	Hindi, Kannada, Malayalam, Tamil and Telugu	F1 score: 97.0% Precision: NA Recall: NA	Supervised	RNN+LSTM
12	Lyu, Chen [115]	Biomedical	Bio Creative II gene mention (GM) corpus & JNLPBA corpus	English	F1 score: 86.6 % Precision:87.9% Recall: 85.3%	Supervised	LSTM-RNN +BLSTM-RNN+CRF
	Part of speech (POS) tagging						
13	Prabha, Jyothsna [116]	Linguistic	Nepali text	Nepali	F1 score:92.0% Precision: 93% Recall:93%	Supervised	RNN+LSTM+GRU
14	Jung and Lee [117]	Linguistic	Korean text	Korean English	F1 score:70.9% Precision:77% Recall:68.0%	Supervised	RNN+LSTM+GRU

6.2 Convolutional Neural Network Model in NLP

Table 5 below represents some recent convolutional neural network approaches for information extraction problems. Convolutional neural network in deep learning has gained explosive prominence in most NLP applications, and Hu, Lu [118] as well as other researchers have established the success and the potential of CNN in NLP applications such as information extraction. CNN is primarily known for extracting the majority of the n-gram attributes from an input. It subsequently formulates the hidden semantic representation

following classification undertakings [110, 119]. While it has been widely used in computer vision and images classification applications, it has also been used as classifiers for some robust text classification tasks, and four or more layers may constitute a typical CNN for text mining.

The table is self-explanatory with particular reference to the information extraction task, the domain application, dataset and the overall performance evaluation of the project.

Table 5.Recent Applications of CNN model in Information Extraction Projects

No	Study & Author(s)	Domain Application	Data set	Language	Performance	Learning Type	Model
	Named Entity Recognition						
15	Khalifa and Shaalan [120]	character-features	Twitter dataset	Arabic	F1 score:79.5% Precision:55.5% Recall:65.3%	Supervised	WC-CNN+CRF
16	Xu, Li [121]	Nested entity relations	Chinese Recognition Corpus	Chinese	F1 Score:95.0%	Supervised	CNN+SVM
17	Ru, Tang [122]	Social networks	Web	English	F1 score:73.7% Precision:76.5% Recall:71.1%	Supervised	CNN

	Text detection and summarisation						
18	Ren, Chen [123]	Text recognition	Ren's, Pan's & Zhou's	Chinese	F1 score:76.0% Precision:81.0% Recall:71.0%	Unsupervised	CNN+TSCD
	Text Comprehension						
19	Qiu, Yoon [124]	Cancer Surveillance	cancer pathology reports	English	Micro-F score:84.2%	Unsupervised	CNN +Networked training
	IE from Images &. Videos						
20	Nagarajan and Perumal [125]	IE from images	Web pages	English	F1 score: Precision: Recall:	Supervised	CNN+ANN
21	Zhang, Lin [126]	Videos	YouTube-8M & Sport-1M	English	F1 score: NA Precision: NA Recall: NA	Supervised	CNN+VWII
	Topic Categorisation						
22	Arras, Horn [127]	Semantic category	20News-groups dataset	English	Acc:79.8%(F=600) Acc:80.2%(F=800) Acc:79.8%(F=600)	Unsupervised	CNN+LRP
	Text Classification						
23	Yoon, Robinson [128]	Cancer Pathology	Cancer pathology reports	English	Micro-F1:77.4% Macro-F1:56.0%	Unsupervised	CNN+ Filter pruning
	Semantic Clustering						
24	Alawad, Yoon [129]	Cancer pathology	Cancer pathology reports	English	Micro-F score:77.5%	Unsupervised	CNN+MTL

7. CONCLUSION AND FUTURE RESEARCH

Given the explosion in social network platforms with their subjective data contents, the urgency to develop computational applications to understand, classify and analyse what humans have written and spoken is critical. Deep learning has become a robust machine learning technique to facilitate these advanced NLP applications. This study reviewed and summarised over 50 deep learning NLP paradigms, according to their respective architecture and task.

Although NLP applications and its related technologies have seen numerous success in recent times, there remain so many challenges that confront the discipline. Natural Language Processing, which is the study of techniques to make computers understand and communicate with humans, is a challenging area. Consequently, the explosion in social media which coincided with these intelligent systems have become a valuable source of data for these applications. Since IE applications are generally domain-specific, the challenges attributable to IE are also domain-oriented. For an instant, a study by [130] Opined, their NLP application outperformed the existing system, however unable to replicate the methods in different domains. Another challenge is topical and semantic categorisation of discourse using social data with regards to varied domains. Furthermore, recent NLP applications seem to be focusing on fact harvesting in various domain. Subsequently, most of this sought-after information is represented in diverse languages, as such, multi-lingual data sources are one of the challenges that confront NLP applications in recent times. Only languages like English, Chinese, Arabic and Nepalis and some Indian languages have largely been used in NLP applications, leaving room for replication of most of the NLP applications to more languages across the globe.

While the NLP applications and experiments reviewed in this study reported the impressive outcome of the various projects, it is obvious that just portions of the results were published as a result of the limitations associated in publishing research outcomes. This review systematically analyses the trajectory of deep learning-NLP and the current trends as well as data used in information extraction applications. One of the prevailing challenges of NLP study, especially in information extraction, has been the absence of huge dataset and their corresponding reusable experimentation methods and tools. This has facilitated the over usage of certain datasets such as Twitter stream and few other large data repositories, particularly in the health sector. Additionally, most information extraction tasks are domain biased. This study is of the view that integrating multiple data sources for an NLP application will greatly improve the performance of any applicable NLP system and multi-domain IE projects.

From the reviewed research materials in this study, portability and scalability remain a current challenge and future area of study in NLP applications. Additionally, the algorithmic pattern of current NLP systems is always subject to change because information extraction is domain specific. Furthermore, concept and context-aware extraction are extremely costly and daunting for a large-scale natural language processing application, although more effort has been dedicated to this area of study.

8. REFERENCES

[1] Bowden, K.K., et al., *Data-driven dialogue systems for social agents*, in *Advanced Social Interaction with Agents*. 2019, Springer. p. 53-56.

[2] Salloum, S.A., et al., *Using text mining techniques for extracting information from research articles*, in *Intelligent natural language processing: Trends and Applications*. 2018, Springer. p. 373-397.

[3] Dash, A., M. Pandey, and S. Rautaray. *Enhanced Entity Extraction Using Big Data Mechanics*. in *International Conference on Advanced Computing Networking and Informatics*. 2019. Springer.

[4] Fianyi, I. and T.A. Zia, *Biometric Technology Solutions to Countering Today's Terrorism*, in *Violent Extremism: Breakthroughs in Research and Practice*. 2019, IGI Global. p. 399-412.

[5] Dong, Z.S., et al., *Social media information sharing for natural disaster response.* Natural Hazards, 2021: p. 1-28.

[6] Jiang, S., et al. *Factoring fact-checks: Structured information extraction from fact-checking articles*. in *Proceedings of The Web Conference 2020*. 2020.

[7] Beel, J., et al. *Research paper recommender system evaluation: a quantitative literature survey*. in *Proceedings of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation*. 2013.

[8] Espadoto, M., et al., *Towards a quantitative survey of dimension reduction techniques.* IEEE Transactions on Visualization and Computer Graphics, 2019.

[9] Thomson, A., M. Kennelly, and K. Toohey, *A systematic quantitative literature review of empirical research on large-scale sport events' social legacies.* Leisure Studies, 2020: p. 1-18.

[10] Vera-Baceta, M.-A., M. Thelwall, and K. Kousha, *Web of Science and Scopus language coverage.* Scientometrics, 2019. **121**(3): p. 1803-1813.

[11] Wang, B., et al., *An overview of climate change vulnerability: a bibliometric analysis based on Web of Science database.* Natural Hazards, 2014. **74**(3): p. 1649-1666.

[12] Sheth, J.N., *How social media will impact marketing media*, in *Social media marketing*. 2018, Springer. p. 3-18.

[13] Xu, W. and Y. Tan, *Semi-supervised Target-oriented Sentiment Classification.* Neurocomputing, 2019.

[14] Maynard, D., K. Bontcheva, and I. Augenstein, *Natural Language Processing for the Semantic Web*, in *Synthesis Lectures on the Semantic Web: Theory and Technology*. 2017, Morgan and Claypool Publishers. p. 1-196.

[15] Lim, S., C.S. Tucker, and S. Kumara, *An unsupervised machine learning model for discovering latent infectious diseases using social media data.* Journal of Biomedical Informatics, 2017. **66**: p. 82-94.

[16] Giannakopoulos, T., et al. *Interactive Text Analysis and Information Extraction*. in *Italian Research Conference on Digital Libraries*. 2019. Springer.

[17] Ertam, F. *Deep learning based text classification with Web Scraping methods*. in *2018 International Conference on Artificial Intelligence and Data Processing, IDAP 2018*. 2019. Institute of Electrical and Electronics Engineers Inc.

[18] Emami, H., H. Shirazi, and A.A. Barforoush, *Web Person Name Disambiguation Using Social Links and Enriched Profile Information.* COMPUTING AND INFORMATICS, 2019. **37**(6): p. 1485-1515.

[19] Singh, A., N. Shukla, and N. Mishra, *Social media data analytics to improve supply chain management in food industries.* Transportation Research Part E: Logistics and Transportation Review, 2018. **114**: p. 398-415.

[20] Kalra, V. and R. Agrawal, *Challenges of Text Analytics in Opinion Mining*, in *Extracting Knowledge From Opinion Mining*. 2019, IGI Global. p. 268-282.

[21] Tarasconi, F., et al. *The role of unstructured data in real-time disaster-related social media monitoring*. in *5th IEEE International Conference on Big Data, Big Data 2017*. 2018. Institute of Electrical and Electronics Engineers Inc.

[22] Jayathilaka, K.M.P.N., A.R. Weerasinghe, and W.M.L.K.N. Wijesekara. *Making sense of large volumes of unstructured email responses*. in *16th International Conference on Advances in ICT for Emerging Regions, ICTer 2016*. 2017. Institute of Electrical and Electronics Engineers Inc.

[23] Musial, K., et al., *Extraction of Multilayered Social Networks from Activity Data.* Scientific World Journal, 2014.

[24] Kang, J., A. Souili, and D. Cavallucci, *Text simplification of patent documents*, in *18th International TRIZ Future Conference on Automated Invention for Smart Industries, TFC 2018*, S. Koziolek, D. Cavallucci, and R. De Guio, Editors. 2018, Springer New York LLC. p. 225-237.

[25] Thomas, A. and S. Sangeetha, *An innovative hybrid approach for extracting named entities from unstructured text data.* Computational Intelligence, 2019.

[26] Alves, T., et al., *Development of text mining tools for information retrieval from patents*, in *11th International Conference on Practical Applications of Computational Biology and Bioinformatics, PACBB 2017*, M. Rocha, et al., Editors. 2017, Springer Verlag. p. 66-73.

[27] Liu, Y. and Y. Huang, *Research on construction of patent dynamic technology effect matrix.* International Journal of Innovative Computing, Information and Control, 2018. **14**(3): p. 1133-1139.

[28] Abilhoa, W.D. and L.N. de Castro, *TKG: A graph-based approach to extract keywords from tweets*, in *11th International Symposium on Distributed Computing and Artificial Intelligence 2014, DCAI 2014*. 2014, Springer Verlag: Salamanca. p. 425-432.

[29] Sato, N. and K. Sato, *Statistical analysis of word usage in biological publications since 1965: Historical delineation highlighting an emergence of function-oriented discourses in contemporary molecular and cellular biology.* Journal of Theoretical Biology, 2019. **462**: p. 293-303.

[30] Gutierrez, F., et al., *A hybrid ontology-based information extraction system.* Journal of Information Science, 2015. **42**(6): p. 798-820.

[31] Ali, T., et al., *Multi-model-based interactive authoring environment for creating shareable medical knowledge.* Computer Methods and Programs in Biomedicine, 2017. **150**: p. 41-72.

[32] Tourassi, G., et al., *The utility of web mining for epidemiological research: Studying the association between parity and cancer risk.* Journal of the American Medical Informatics Association, 2016. **23**(3): p. 588-595.

[33] Jebbor, F. and L. Benhlima, *Ontology-based enhancement of rule learning for information extraction.* Journal of Theoretical and Applied Information Technology, 2018. **96**(23): p. 7876-7891.

[34] Saad, S. and M.K. Mansor, *Named entity recognition approach for malay crime news retrieval.* GEMA Online Journal of Language Studies, 2018. **18**(4): p. 216-235.

[35] Klein, A., et al. *A Combined Approach for Extracting Financial Instrument-Specific Investor Sentiment from Weblogs*. in *Wirtschaftsinformatik*. 2013.

[36] Abrantes, D. and J. Cordeiro. *Extracting Adverse Drug Effects from User Experiences: A Baseline*. in *31st IEEE International Symposium on Computer-Based Medical Systems, CBMS 2018*. 2018. Institute of Electrical and Electronics Engineers Inc.

[37] Timonin, A.Y., A.S. Bozhday, and A.M. Bershadsky. *Analysis of unstructured text data for a person social profile*. in *2017 International Conference on Electronic Governance and Open Society, eGose 2017*. 2017. Association for Computing Machinery.

[38] Tarasconi, F., et al., *The Role of Unstructured Data in Real-Time Disaster-related Social Media Monitoring*, in *2017 Ieee International Conference on Big Data*, J.Y. Nie, et al., Editors. 2017. p. 3769-3778.

[39] Bin Tareaf, R., et al., *Personality Exploration System for Online Social Networks: Facebook Brands As a Use Case*. 2018 Ieee/Wic/Acm International Conference on Web Intelligence. 2018. 301-309.

[40] Abdelsadek, Y., et al., *Community extraction and visualization in social networks applied to Twitter.* Information Sciences, 2018. **424**: p. 204-223.

[41] Beigi, G., et al., *An overview of sentiment analysis in social media and its applications in disaster relief*, in *Sentiment analysis and ontology engineering*. 2016, Springer. p. 313-340.

[42] Pandey, N. and S. Natarajan, *How Social Media can Contribute during Disaster Events?* 2016 International Conference on Advances in Computing, Communications and Informatics, ed. J. Wu, et al. 2016. 1352-1356.

[43] Luong, T.L., et al. *Intent extraction from social media texts using sequential segmentation and deep learning models*. in *9th International Conference on Knowledge and Systems Engineering, KSE 2017*. 2017. Institute of Electrical and Electronics Engineers Inc.

[44] Ramírez-Cifuentes, D., M. Mayans, and A. Freire, *Early risk detection of anorexia on social media*, in *5th International Conference on Internet Science, INSCI 2018*, S.S. Bodrunova, Editor. 2018, Springer Verlag. p. 3-14.

[45] Nikfarjam, A., et al., *Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features.* Journal of the American Medical Informatics Association, 2015. **22**(3): p. 671-681.

[46] Yang, T.F., et al., *Social Media Big Data Mining and Spatio-Temporal Analysis on Public Emotions for Disaster Mitigation.* Isprs International Journal of Geo-Information, 2019. **8**(1).

[47] Sarker, A., A. Nikfarjam, and G. Gonzalez, *Social media mining shared task workshop.* 21st Pacific Symposium on Biocomputing, PSB 2016, 2016: p. 581-592.

[48] Aung, S.S. and M.S. Wai, *Domain independent feature extraction using rule based approach.* Advances in Science, Technology and Engineering Systems, 2018. **3**(1): p. 218-224.

[49] Žižka, J. and F. Dařena. *Revealing potential changes of significant terms in streams of textual data written in natural languages using windowing and text mining*. in *Artificial Intelligence and Natural Language and Information Extraction, Social Media and Web Search FRUCT Conference, AINL-ISMW FRUCT 2015*. 2016. Institute of Electrical and Electronics Engineers Inc.

[50] Sparks, K.A., et al. *Facility detection and popularity assessment from text classification of social media and crowdsourced data*. in *10th Workshop on Geographic Information Retrieval, GIR 2016*. 2016. Association for Computing Machinery, Inc.

[51] Sang, L., et al., *WEFEST: Word Embedding Feature Extension for Short Text Classification*, in *2016 Ieee 16th International Conference on Data Mining Workshops*, C. Domeniconi, et al., Editors. 2016. p. 677-683.

[52] Fianyi, I. and T. Zia, *Biometric technology solutions to countering today's terrorism*, in *Violent Extremism: Breakthroughs in Research and Practice*. 2018, IGI Global. p. 399-412.

[53] Papadimitriou, D., et al., *Goals in Social Media, Information Retrieval and Intelligent Agents*, in *2015 Ieee 31st International Conference on Data Engineering.* 2015. p. 1538-1540.

[54] Onal, K.D. and P. Karagoz. *Named entity recognition*

from scratch on social media. in *6th International Workshop on Mining Ubiquitous and Social Environments, MUSE 2015*. 2015. CEUR-WS.

[55] Fabian, B., A. Baumann, and M. Keil. *Privacy on reddit? Towards large-scale user classification*. in *23rd European Conference on Information Systems, ECIS 2015*. 2015. Association for Information Systems.

[56] Zolnoori, M., et al., *A systematic approach for developing a corpus of patient reported adverse drug events: A case study for SSRI and SNRI medications.* Journal of Biomedical Informatics, 2019. **90**.

[57] Saad Missen, M.M., et al., *SentiML ++: an extension of the SentiML sentiment annotation scheme.* New Review of Hypermedia and Multimedia, 2018. **24**(1): p. 28-43.

[58] Funk, C., et al. *Community-driven crowdsourcing: Data collection with local developers*. in *11th International Conference on Language Resources and Evaluation, LREC 2018*. 2019. European Language Resources Association (ELRA).

[59] Chen, Y.F., et al., *Review on Rumor Detection of Online Social Networks.* Jisuanji Xuebao/Chinese Journal of Computers, 2018. **41**(7): p. 1648-1677.

[60] Basile, V., M. Lai, and M. Sanguinetti. *Long-term social media data collection at the University of Turin*. in *5th Italian Conference on Computational Linguistics, CLiC-it 2018*. 2018. CEUR-WS.

[61] Abid, A., et al. *An extraction and unification methodology for social networks data: An application to public security*. in *19th International Conference on Information Integration and Web-Based Applications and Services, iiWAS2017*. 2017. Association for Computing Machinery.

[62] Zaghouani, W. and A. Charfi. *AraP-Tweet: A large multi-dialect twitter corpus for gender, age and language variety identification*. in *11th International Conference on Language Resources and Evaluation, LREC 2018*. 2019. European Language Resources Association (ELRA).

[63] Kuwertz, A., et al., *Applying knowledge-based reasoning for information fusion in intelligence, surveillance, and reconnaissance*, in *13th IEEE International Conference on Multisensor Integration and Fusion, IEEE MFI 2017*, H. Ko, S. Lee, and S. Oh, Editors. 2018, Springer Verlag. p. 119-139.

[64] Farnadi, G., et al. *User profiling through deep multimodal fusion*. in *11th ACM International Conference on Web Search and Data Mining, WSDM 2018*. 2018. Association for Computing Machinery, Inc.

[65] Zhang, S., et al., *Deep learning based recommender system: A survey and new perspectives.* ACM Computing Surveys (CSUR), 2019. **52**(1): p. 5.

[66] Liu, T., et al., *Deep learning-based super-resolution in coherent imaging systems.* Scientific Reports, 2019. **9**(1).

[67] Zhou, K., et al., *Deep sentiment hashing for text retrieval in social CIoT.* Future Generation Computer Systems, 2018. **86**: p. 362-371.

[68] Zhang, X., et al., *Exploring Deep Recurrent Convolution Neural Networks for Subjectivity Classification.* IEEE Access, 2019. **7**: p. 347-357.

[69] Liu, B., *Sentiment analysis: Mining opinions, sentiments, and emotions*. 2015: Cambridge University Press.

[70] Mao, W., W. Feng, and X. Liang, *A novel deep output kernel learning method for bearing fault structural diagnosis.* Mechanical Systems and Signal Processing, 2019. **117**: p. 293-318.

[71] Coates, A., A. Ng, and H. Lee. *An analysis of single-layer networks in unsupervised feature learning*. in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. 2011.

[72] Marcus, G., *Deep learning: A critical appraisal.* arXiv preprint arXiv:1801.00631, 2018.

[73] Bengio, Y., *Learning deep architectures for AI.* Foundations and trends® in Machine Learning, 2009. **2**(1): p. 1-127.

[74] Scherer, R., *Feature detection*, in *Studies in Computational Intelligence*. 2020, Springer Verlag. p. 7-32.

[75] Kamruzzaman, A., Y. Alhwaiti, and C.C. Tappert, *Developing a deep learning model to implement rosenblatt's experiential memory brain model*, in *Lecture Notes in Networks and Systems*. 2020, Springer. p. 248-261.

[76] Zhao, Z., et al. *Deep neural network based protein-protein interaction extraction from biomedical literature*. in *IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2015*. 2015. Institute of Electrical and Electronics Engineers Inc.

[77] Agnati, L.F., et al., *The brain as a "hyper-network": the key role of neural networks as main producers of the integrated brain actions especially via the "broadcasted" neuroconnectomics.* Journal of Neural Transmission, 2018. **125**(6): p. 1-15.

[78] Qi, Y., et al., *Deep learning for character-based information extraction*, in *36th European Conference on Information Retrieval, ECIR 2014*. 2014, Springer Verlag: Amsterdam. p. 668-674.

[79] Jin, Y., D. Wu, and W. Guo, *Attention-Based LSTM with Filter Mechanism for Entity Relation Classification.* Symmetry, 2020. **12**(10): p. 1729.

[80] Liu, G. and J. Guo, *Bidirectional LSTM with attention mechanism and convolutional layer for text classification.* Neurocomputing, 2019. **337**: p. 325-338.

[81] Li, W., D. Guo, and X. Fang, *Multimodal architecture for video captioning with memory networks and an attention mechanism.* Pattern Recognition Letters, 2018. **105**: p. 23-29.

[82] Wu, X., et al., *Conditional BERT Contextual Augmentation*, J.M.F. Rodrigues, et al., Editors. 2019, Springer Verlag. p. 84-95.

[83] Zirikly, A. and M. Diab, *ANEAR: Automatic named entity aliasing resolution*, in *18th International Conference on Application of Natural Language to Information Systems, NLDB 2013*. 2013: Salford. p. 213-224.

[84] Zhou, X., et al. *Fake News: Fundamental theories, detection strategies and challenges*. in *12th ACM International Conference on Web Search and Data*

Mining, WSDM 2019. 2019. Association for Computing Machinery, Inc.

[85] Zhang, Y., et al., *Empower event detection with bi-directional neural language model.* Knowledge-Based Systems, 2019. **167**: p. 87-97.

[86] Zaw, M. and P. Tandayya. *Multi-level Sentiment Information Extraction Using the CRbSA Algorithm*. in *15th International Joint Conference on Computer Science and Software Engineering, JCSSE 2018*. 2018. Institute of Electrical and Electronics Engineers Inc.

[87] Zahra, K., et al. *A framework for user characterization based on tweets using machine learning algorithms*. in *7th International Conference on Network, Communication and Computing, ICNCC 2018*. 2018. Association for Computing Machinery.

[88] Yu, Z., et al., *Improving the utility of MeSH® terms using the TopicalMeSH representation.* Journal of Biomedical Informatics, 2016. **61**: p. 77-86.

[89] Yin, Z., et al., *A scalable framework to detect personal health mentions on Twitter.* Journal of Medical Internet Research, 2015. **17**(6): p. e138.

[90] Yang, Q., et al., *Building enhanced link context by logical sitemap*, in *6th International Conference on Knowledge Science, Engineering and Management, KSEM 2013*. 2013: Dalian. p. 36-47.

[91] Xu, A., et al. *A new chatbot for customer service on social media*. in *2017 ACM SIGCHI Conference on Human Factors in Computing Systems, CHI 2017*. 2017. Association for Computing Machinery.

[92] Wu, Y., et al., *Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots.* arXiv preprint arXiv:1612.01627, 2016.

[93] Weitzel, L., R.C. Prati, and R.F. Aguiar, *The comprehension of figurative language: What is the influence of irony and sarcasm on NLP techniques?*, in *Studies in Computational Intelligence*. 2016, Springer Verlag. p. 49-74.

[94] Wang, W., et al. *A multiple instance learning framework for identifying key sentences and detecting events*. in *25th ACM International Conference on Information and Knowledge Management, CIKM 2016*. 2016. Association for Computing Machinery.

[95] Trusov, R., et al. *Multi-representation approach to text regression of financial risks*. in *Artificial Intelligence and Natural Language and Information Extraction, Social Media and Web Search FRUCT Conference, AINL-ISMW FRUCT 2015*. 2016. Institute of Electrical and Electronics Engineers Inc.

[96] Tian, T. and W. Xu, *Chinese lexical normalization based on information extraction: An experimental study*, in *26th International Conference on Artificial Neural Networks, ICANN 2017*, A. Lintas, et al., Editors. 2017, Springer Verlag. p. 216-223.

[97] Tian, B. and C. Xing, *Deep Learning Based Temporal Information Extraction Framework on Chinese Electronic Health Records*, X. Wang, et al., Editors. 2018, Springer Verlag. p. 203-214.

[98] Thomas, P., et al. *Streaming text analytics for real-time event recognition*. in *11th International Conference on Recent Advances in Natural Language Processing, RANLP 2017*. 2017. Association for Computational Linguistics (ACL).

[99] Srivastava, S.K., S.K. Singh, and J.S. Suri, *Effect of incremental feature enrichment on healthcare text classification system: A machine learning paradigm.* Computer Methods and Programs in Biomedicine, 2019. **172**: p. 35-51.

[100] Sohangir, S. and D. Wang. *Finding expert authors in financial forum using deep learning methods*. in *2nd IEEE International Conference on Robotic Computing, IRC 2018*. 2018. Institute of Electrical and Electronics Engineers Inc.

[101] Li, L., L. Jin, and D. Huang, *Exploring recurrent neural networks to detect named entities from biomedical text*, M. Sun, et al., Editors. 2015, Springer Verlag. p. 279-290.

[102] Wang, W.Y., J. Li, and X. He. *Deep reinforcement learning for NLP*. in *56th Annual Meeting of the Association for Computational Linguistics, ACL 2018*. 2018. Association for Computational Linguistics (ACL).

[103] Goldberg, Y., *Neural network methods for natural language processing.* Computational Linguistics, 2018. **44**(1): p. 194-195.

[104] Viani, N., et al., *Supervised methods to extract clinical events from cardiology reports in Italian.* Journal of Biomedical Informatics, 2019. **95**.

[105] Gupta, S., et al., *Semi-Supervised Recurrent Neural Network for Adverse Drug Reaction mention extraction.* BMC Bioinformatics, 2018. **19**.

[106] Gao, S., et al., *Hierarchical attention networks for information extraction from cancer pathology reports.* Journal of the American Medical Informatics Association, 2018. **25**(3): p. 321-330.

[107] Xiang, X., et al., *Skeleton to abstraction: An attentive information extraction schema for enhancing the saliency of text summarization.* Information (Switzerland), 2018. **9**(9).

[108] Liu, S., Y. Li, and B. Fan, *Hierarchical RNN for few-shot information extraction learning*, Q. Zhou, et al., Editors. 2018, Springer Verlag. p. 227-239.

[109] Yin, Z., K.H. Chang, and R. Zhang. *DeepProbe: Information directed sequence understanding and chatbot design via recurrent neural networks*. 2017. Association for Computing Machinery.

[110] Banerjee, I., et al., *Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification.* Artificial Intelligence in Medicine, 2019. **97**: p. 79-88.

[111] Banik, N. and M.H.H. Rahman. *GRU based Named Entity Recognition System for Bangla Online Newspapers*. 2019. Institute of Electrical and Electronics Engineers Inc.

[112] Zhang, C., et al., *Multi-Gram CNN-Based Self-Attention Model for Relation Classification.* IEEE Access, 2019. **7**: p. 5343-5357.

[113] Chowdhury, S., et al., *A multitask bi-directional RNN model for named entity recognition on Chinese electronic medical records.* BMC Bioinformatics, 2018. **19**.

[114] Nagesh Bhattu, S., N. Satya Krishna, and D.V.L.N. Somayajulu. *Idrbt-team-a@IECSIL-FIRE-2018: Named Entity Recognition of Indian languages using Bi-LSTM*. 2018. CEUR-WS.

[115] Lyu, C., et al., *Long short-term memory RNN for biomedical named entity recognition.* BMC Bioinformatics, 2017. **18**(1).

[116] Prabha, G., et al. *A Deep Learning Approach for Part-of-Speech Tagging in Nepali Language*. 2018. Institute of Electrical and Electronics Engineers Inc.

[117] Jung, S. and C. Lee, *Deep Neural Architecture for Recovering Dropped Pronouns in Korean.* ETRI Journal, 2018. **40**(2): p. 257-265.

[118] Hu, B., et al. *Convolutional neural network architectures for matching natural language sentences*. in *Advances in neural information processing systems*. 2014.

[119] Chen, W.F. and L.W. Ku. *UTCNN: A deep learning model of stance classification on social media text*. in *26th International Conference on Computational Linguistics, COLING 2016*. 2016. Association for Computational Linguistics, ACL Anthology.

[120] Khalifa, M. and K. Shaalan, *Character convolutions for Arabic Named Entity Recognition with Long Short-Term Memory Networks.* Computer Speech and Language, 2019. **58**: p. 335-346.

[121] Xu, H., et al., *Research on Chinese Nested Named Entity Relation Extraction.* Beijing Daxue Xuebao (Ziran Kexue Ban)/Acta Scientiarum Naturalium Universitatis Pekinensis, 2019. **55**(1): p. 8-14.

[122] Ru, C., et al. *Syntactic representation learning for open information extraction on web*. 2019. International World Wide Web Conferences Steering Committee.

[123] Ren, X., et al. *A novel text structure feature extractor for Chinese scene text detection and recognition*. 2017. Institute of Electrical and Electronics Engineers Inc.

[124] Qiu, J.X., et al., *Scalable deep text comprehension for Cancer surveillance on high-performance computing.* BMC Bioinformatics, 2018. **19**.

[125] Nagarajan, S. and K. Perumal. *A deep neural network for information extraction from web pages*. 2018. Institute of Electrical and Electronics Engineers Inc.

[126] Zhang, C., et al., *CNN-VWII: An efficient approach for large-scale video retrieval by image queries.* Pattern Recognition Letters, 2019. **123**: p. 82-88.

[127] Arras, L., et al., *"What is relevant in a text document?": An interpretable machine learning approach.* PLoS ONE, 2017. **12**(8).

[128] Yoon, H.J., et al. *Filter pruning of Convolutional Neural Networks for text classification: A case study of cancer pathology report comprehension*. 2018. Institute of Electrical and Electronics Engineers Inc.

[129] Alawad, M., H.J. Yoon, and G.D. Tourassi. *Coarse-to-fine multi-task training of convolutional neural networks for automated information extraction from cancer pathology reports*. 2018. Institute of Electrical and Electronics Engineers Inc.

[130] Conway, L.G., K.R. Conway, and S.C. Houck, *Validating automated integrative complexity: Natural language processing and the Donald Trump test.* Journal of Social and Political Psychology, 2020. **8**(2): p. 504-524.