EasyChair Preprint

№ 8240

Systematic Review on Uses of NLP

Awias Sarosh Ansari Ansari, Gaurav Bharat Wagh Wagh,
Kaushal Jayant Lohar Lohar and
Shaikh Sohail Ahmed Kaleem Ahmed Shaikh

June 10, 2022

Systematic Review on Uses of NLP

Awais Sarosh Ansari
Department of Mechanical Engineering
Shri Vile Parle Kelavani Mandal's institute of Technology Dhule.
awaisahmad.ansari@svkmmumbai.onmicrosoft.com

Gaurav Bharat Wagh
Department of Mechanical Engineering
Shri Vile Parle Kelavani Mandal's institute of Technology Dhule.
gaurav.wagh@svkmmumbai.onmicrosoft.com

Kaushal Jayant Lohar
Department of Mechanical Engineering
Shri Vile Parle Kelavani Mandal's institute of Technology Dhule.
Kaushal.lohar@svkmmumbai.onmicrosoft.com

Sohail Kaleem Shaikh
Department of Mechanical Engineering
Shri Vile Parle Kelavani Mandal's institute of Technology Dhule.
Sohailahmed.shaikh@svkmmumbai.onmicrosoft.com

Abstract

Natural Language Processing (NLP) is a subfield of Artificial Intelligence that is receiving a lot of attention in terms of research and development as a result of the growing number of applications. Natural Language Processing (NLP) is a computerised method of analysing texts. NLP entails obtaining information on how humans comprehend and use language. This is being done in order to build tools and approaches that will allow computers to interpret and manipulate natural languages in order to execute various tasks. Natural language processing is being used in a variety of industries, including healthcare, finance, manufacturing, and education, retail, and customer service.
This paper examines the literature on the uses of Natural Language Processing (NLP) in machine translation, text analysis, and conversational & question answering systems. It also includes a brief history of NLP as well as earlier NLP research. It is based on the examination of documents. This research article may be useful to anyone interested in studying and learning about natural language processing (NLP) and its uses.

Keywords

Natural Language Processing, Question Answering Systems, Machine Translation, Text Analysis, Syntactic and semantic analysis

Objectives

This paper summarizes information about NLP, including a brief overview, brief history, and past studies, based on document analysis. It then focuses on application of NLP in Machine Translation, Text Analysis, Conversational & Question Answering Systems. The review paper is meant to provide scholarly colleagues and students, particularly freshmen, with an understanding of the usage and applications of NLP.

Introduction

Natural Language Processing is a branch of Linguistics, Computer Science, and Artificial Intelligence concerned with Human-Computer-Interaction, particularly how to design computers to process and evaluate huge volumes of Natural Language Data. Natural Language Processing (NLP) is concerned with the use of computational methods to automatically represent and interpret many types of Human (Natural) Language inputs and communicate using Human-Computer-Interfaces (HCI). "Computational Linguistics" is another term for it.

NLP is a type of Artificial Intelligence that assists computers in understanding, manipulating, and interpreting human language. Natural language processing encompasses a wide range of topics, including computer science and computational linguistics, with the goal of closing the gap between human machine communication and machine comprehension. The complexity and diversity of human languages is astounding. We've expressed ourselves in many different ways, both verbally and in writing. [1]

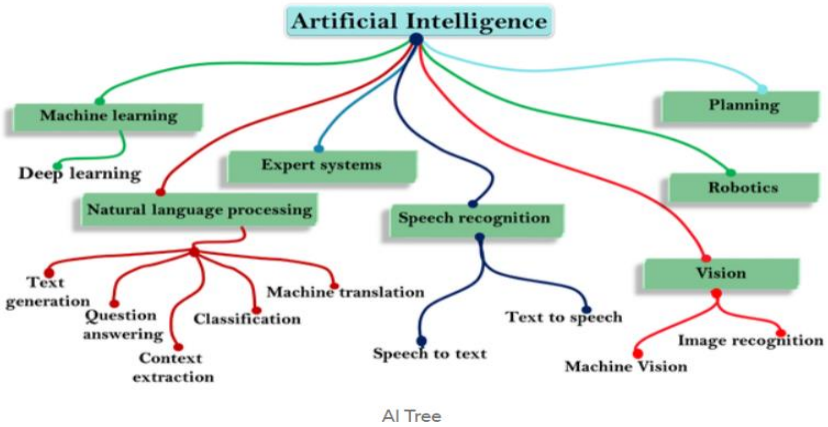


Fig.1 NLP in A.I (Mohadikar)

Natural Language Processing is important in Artificial Intelligence because it aids in the resolution of ambiguity in the language and the addition of a much more helpful numerical structure for the data to words in a variety of downstream applications such as speech recognition and text analytics. Natural Language Processing (NLP) is the study of how computers comprehend and translate human language. NLP allows computers to comprehend written or spoken language and perform tasks such as translation, keyword extraction, subject categorization, and more [2].

Natural Language Processing (NLP) is a field of study that investigates how computers can interpret and modify natural language text or speech to perform useful tasks, according to various academics. ([3]; [4])

NLP is a theoretically motivated set of computational approaches for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis in order to achieve human-like language processing for a variety of tasks or applications [7]. The term "Natural Language Processing" (NLP) refers to the function of software or hardware components in a computer system that analyses or synthesize spoken or written language [5]. Language is more than just a means of communication. Language is a set of tools that allows us to convey meanings, but it is not best understood of as a technique of "encoding" meanings [6]. Lexical (structure) analysis, parsing, semantic analysis, discourse integration, and pragmatic analysis are some of the steps in natural language processing.

Brief History

Although work on NLP extends back to roughly 1950 with the establishment of the "Turing Test" and a rule-based syntactic structure system in 1957. Due to limited processing capacity, progress was slow until 1990, and systems relied on complicated sets of handwritten rules and a limited vocabulary. With the advent of machine learning and the constant rise in processing capacity, interest in research and applications has risen recently. Speech recognition, language processing, dialogue systems, and the application of deep learning techniques are some of the most recent major breakthroughs in NLP. While there are still many obstacles in NLP (such as human-computer interfaces), it has sparked a lot of academic interest and opened up a lot of possibilities for employing the techniques in robotics, automation, and digital transformation.

Research Work of NLP

Until 1990, the majority of research focused on natural language processing (NLP) and machine translation. Recent NLP research has tapped into the potential of statistical models, machine learning, and deep learning technologies, all of which are based on a data-driven approach. Some artificial intelligence and Deep Learning research subjects intersect with Natural Language Processing research issues. These methods have recently become popular for performing NLP jobs in the most effective manner. The ACL 2018 Main Conference invited papers in 21areas which are Dialogue and Interactive Systems, Discourse and Pragmatics, Document Analysis, Generation, Information Extraction and Text Mining, Linguistic Theories, Cognitive Modeling and Psycholinguistics, Machine Learning, Machine Translation, Multidisciplinary, Multilingualism, Phonology, Morphology and Word Segmentation, Question Answering, Resources and Evaluation, Sentence-level Semantics, Sentiment Analysis and Argument Mining, Social Media, Summarization, Tagging, Chunking, Syntax and Parsing, Textual Inference and Other Areas of Semantics, Vision, Robotics, Multimodal, Grounding and Speech, Word-level Semantics. On the above topic, the conference accepted 258 long papers and 156 short papers ([8];[9]).

NLP Techniques

Natural Language Processing (NLP) employs two techniques to help computers comprehend text:
Syntactic analytics and semantic analysis.

Syntactic Analysis

Syntactic analysis, also known as parsing, evaluates text using basic grammatical principles to recognize sentence structure, organize words, and connect them.
Some of its principal subtasks are:

• Tokenization is the process of breaking down a text into smaller chunks called tokens to make material handling easier.

• The part of speech tag identifies tokens such as verb, adverb, adjective, substance, and so on. This aids in the interpretation of words (for example, the term "book" refers to several objects whether employed as a verb or a substantive).

• Lemmatization and stemming is the process of reducing inflected phrases to their simplest form to make analysis easier.

• Stop-word removal frequently eliminates words with no semitone value, such as me, them, have, and so on [10].

Semantic Analysis

Text capture is required for semantic analysis. It begins by examining the meaning of each phrase (lexical semantics). The order of the words, as well as what they mean, is then reviewed in context. The following are the primary tasks of semantical analysis: The goal of word meaning disambiguation is to figure out what a word means in a specific context [11].

Uses and Implementation of NLP

NLP aspires to take over human-machine interaction to the point that conversing with a machine is as simple as conversing with a human. NLP is still used to organise unstructured data and make it understandable to machines. Natural language text processing and summarization, machine translation, user interfaces, multilingual and cross-language information retrieval, speech recognition, artificial intelligence and expert systems, and so on are all examples of NLP applications ([12]; [13]). The fields where NLP is being applied and used are as follows:

- Conversational System
- Question Answering System
- Text Analytics
- Machine Translation
- Healthcare
- Automotive
- Finance
- Manufacturing
- Retail
- Education

Industrial applications of NLP can be broadly classified into 3 categories:
Conversational systems, Text Analytics, Machine translation
Now in this paper, we will focus and review the literature on these 3 categories:

1. Conversational & Question Answering Systems
2. Text Analysis
3. Machine Translation

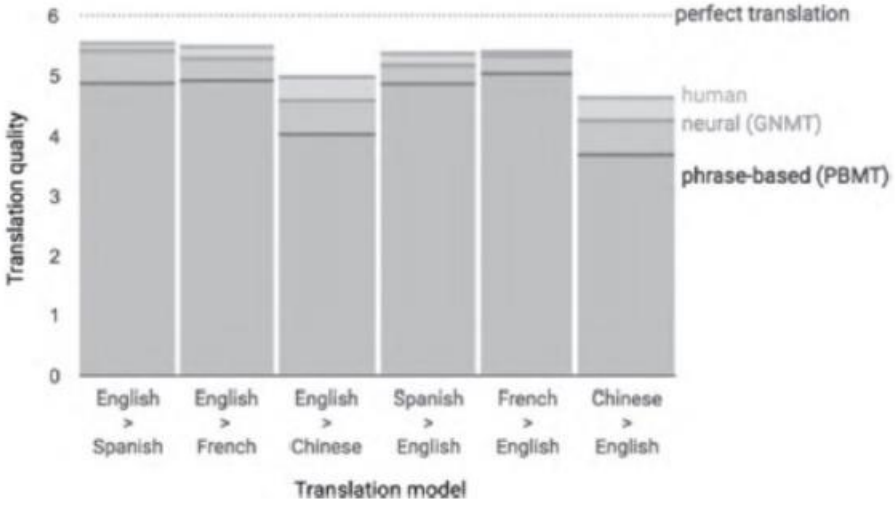
1. Machine Translation

Introduction

The job of automatically translating one natural language into another while preserving the meaning of the input text is known as machine translation [14]. Machine translation relies on natural language processing. Its development process is almost identical to that of machine translation in the past. Most popular application of machine translation is Google translator. Other machine translation software's are also used in speech translation and teaching.

Literature Review

On September 27, 2016, Google Brain Group scientists Quoc V. Le and Mike Schuster published a blog post on Google Research Blog titled "A product-scale neural network for machine translation," claiming that Google had made another major breakthrough in the field of machine translation by introducing a new machine translation system called GNMT, ten years after launching the phrase-based machine translation system "Google Translate." (Simon and Schuster, 2016). Google also issued a paper on arrival that explained the technical workings of the GNMT (Wu et al. 2016). The traditional phrase-based statistical machine translation model GNMT considerably lowered the translation error rate between numerous important language pairs by 55 percent to 85 percent, according to experimental results using Wikipedia and news corpora as test data. Figure 1 indicates that machine translation quality from French to English and English to Spanish is extremely close to human translation quality [15].



Machine Translation and Natural Language Processing in Statistics Corpus

Machine translation systems might be classified into three categories prior to 1990: direct translation, intermediate language translation, and transfer translation (Hutchins, 2009: 505-509). In his special report to the Fourth High-Level Meeting on Machine Translation, held in Kobe, Japan in July 1993, the eminent British researcher W.J. Hutchins declared that the advancement of machine translation had entered a new age since 1989. The introduction of a corpus in machine translation technology, specifically the corpus turned database into machine translation by means of natural

language processing [16], is a significant symbol of the new era. Corpus-based machine translation systems have advanced quickly in recent years, achieving amazing results. Warren Weaver proposed his approach of machine translation utilizing the concept of decoding passwords in his memorandum "Translation" as early as 1947. This so-called "password interpretation" method is simply natural language processing. He intended to overcome machine translation challenges using a statistics-based strategy [17]. IBM's Peter Brown and other scientists developed mathematical models for statistical machine translation based on Weaver's theories. As demonstrated in Figure 1, statistical-based machine translation considers machine translation difficulties as a noisy channel problem.

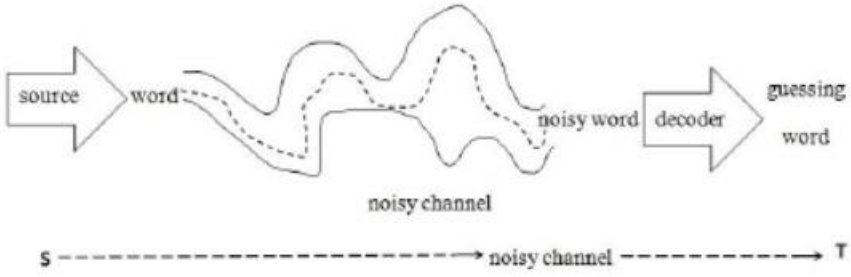


Fig 1. noisy channel model

This is how machine translation works: Source language is distorted as it passes through a noisy channel and appears as Target language at the other end. The challenge with translation is determining the most likely Source language from the observed Target language. In the channel sense, the Source language is the input and the Target language in the sense of translation, whereas in the channel sense, the Target language is the output and the Source language in the sense of translation. From this perspective, any statement in one language could be a translation of certain words in another, but the possibilities of these sentences differ. The goal of machine translation is to locate the sentence with the highest probability, i.e., to calculate the most likely translation of the source language T into all potential target languages S. Stack search can be employed due to the vast number of S. The table structure is the most important data structure in stack search, as it holds the most promising S matching to T. The algorithm keeps looping, extending some of the most promising outcomes with each cycle until the table contains an "S" with a much higher score than the other results. Stack search is a bad approach since it can lead to erroneous translations and so does not guarantee optimal outcomes [18].

The statistical machine translation system's task, as can be seen, is to discover the sentence with the highest probability as a translation result in all potential target languages. The Bayes formula can be used to calculate the probability value.

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Where;
T = target language
S = source language
P(T) = language model of the target language
P(S|T) = translation model of S under given T
Since the denominator P(S) on the right side of the equation is independent of T, finding the maximum value of P(T|S) is equivalent to finding a T so that the multiplication P(T)P(S|T) of the numerator on the right side of the equation is the maximum, that is:

$$T = argmax\ P(T)P(S|T)$$

The process of solving the target language sentence T closest to the real target in the supplied source language sentence S is equivalent to the decoding process in the noisy channel model, according to the language and translation models. Braun and other IBM researchers developed Candide, an English-French machine translation system based on statistical machine translation and using multilingual British and French legislative discussion records as a bilingual corpus.

Table 1 Comparison of Systran & Candide

	Fluency		Faithfulness		Time Ratio	
	1992	1993	1992	1993	1992	1993
Systran	.466	.540	.686	.743		
Candide	.511	.580	.575	.670		
Transman	.819	.838	.837	.850	.688	.625
Manual		.833		.840		

Table 1 shows the results of ARPA's (Advanced Research Projects Agency) tests on a variety of machine translation systems. The first line shows the outcome of the famous Syzran method, the second line shows Candide's translation, the third line shows Candide's manual proofreading, and the fourth line shows a totally human translation. Fluency and Adequacy are the two evaluation indicators. IBM developed Transman, a post-translation modification tool. The Time Ratio displays the proportion of time spent on proofreading by Candide plus Transman vs. time spent on pure manual translation. Candide's natural language processing system has outperformed the traditional Syzran system in terms of index.

Neural Machine Translation & Natural Language Processing

End-to-end neutral machine translation has progressed rapidly since 2014. The translation quality has improved dramatically when compared to statistical machine translation. In 30 languages, Figure 2 compares the experimental outcomes of statistical machine translation versus neural machine translation [19]. In 27 of the natural language processing tasks, neural machine translation outperforms statistical machine translation. As a result, commercial online machine translation systems such as Google, Microsoft, Baidu, and Segou now use neural machine translation instead of statistical machine translation as the fundamental technology.

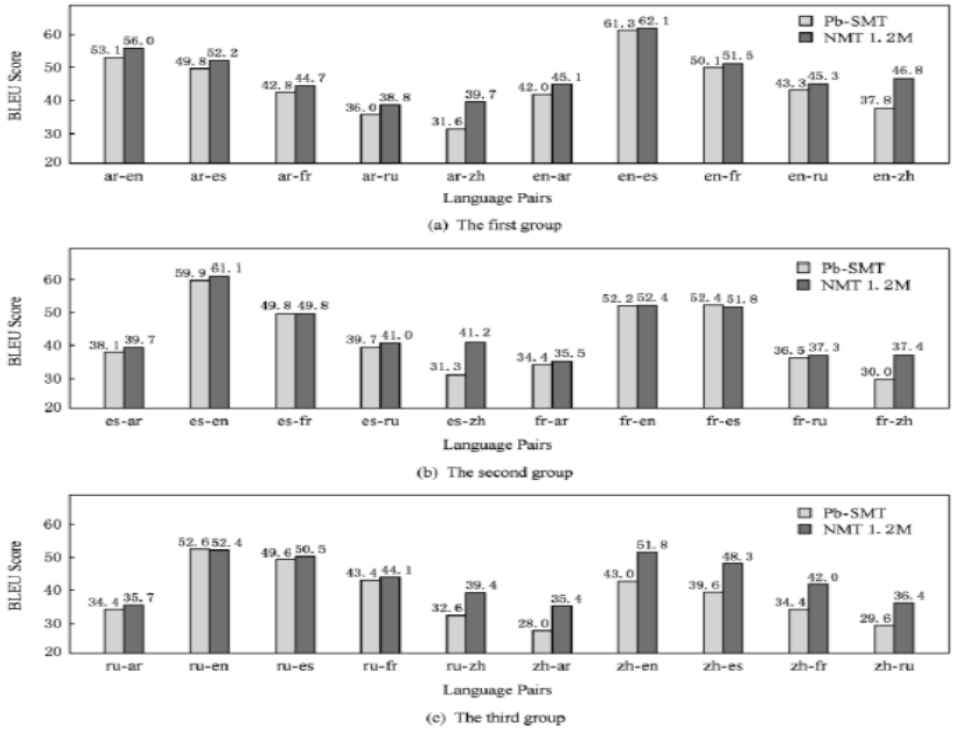


Fig 2. Comparison between statistical machine translation and neural machine translation on 30 Language pairs.

The core idea behind end-to-end neural machine translation is to use neural networks to provide autonomous translation across natural languages. As a result, in order to convert sequences to sequences, neural machine translation often employs an encoder-decoder framework [20]. Taking Figure 3 as an example, given a Chinese sentence "布什 与 沙龙 举行 了 会谈," the encoder-decoder framework first generates a vector representation for each Chinese word and then goes from left to right through a recurrent neural network generating a vector representation of the entire Chinese sentence. Among them, "</s>" indicates the ending terminator. The source language's recursive neural network is referred to as an encoder because it converts the source language sentence into a dense, continuous real number vector.

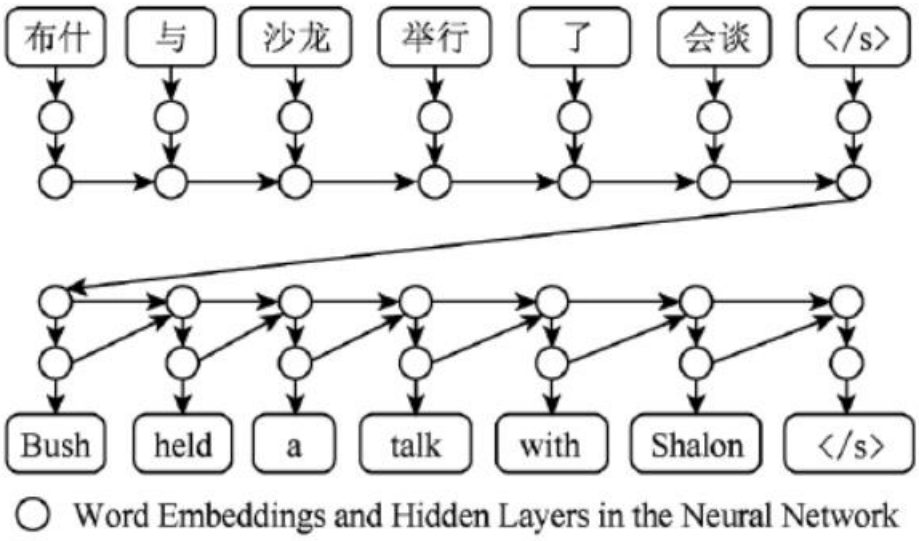
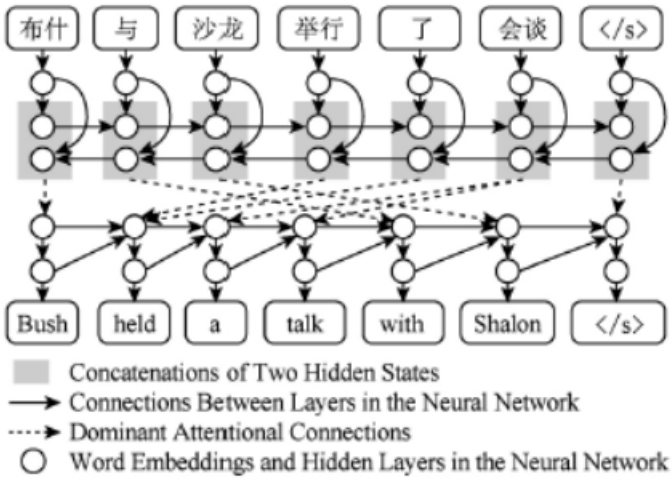


Figure 3 The Encoder-decoder Framework

After that, the target language uses another recursive neural network to reversely decode the source language sentence vector to generate the English sentence "Bush held a talk with Shalon </s>". The entire decoding process is generated word by word, and when the end of the sentence "</s>" is generated, the decoding process terminates. A decoder is the recurrent neural network employed by the target language. It's worth noting that each newly created English word is used as historical data to create the following English term. As a result, the decoder may be thought of as a language model that contains the target language information from the source language. The purpose of neural machine translation based on the attention mechanism is to construct a vector representation containing global information for each source language word, rather than a vector representation for the full source language phrase.



The advantage of this method is that each source language word's vector representation incorporates both left and right context information. The decoder dynamically determines the source language context associated with each target language word on the target language side. For example, when the English word "Bush" is generated, the Chinese word "布什" is the most relevant, and the words "Hold" and "Last" may not be relevant. It is only necessary to pass the "Bush" vector representation as the source context to the target. When the English word "held" is generated, the most relevant Chinese words are "举行" and "了." Therefore, the attention mechanism alters the way information is transmitted and can dynamically determine the most relevant context, allowing it to better tackle long-distance information transmission challenges and increase neural machine translation performance dramatically. As a result, the encoder-decoder model based on the attention mechanism has become the most extensively utilized way of neural machine translation.

Conclusion

Natural language processing is a prerequisite and condition for machine translation. In terms of natural language processing, neural machine translation not only has more generality, but also reflects the power of big data and big data thinking. In addition, neural machine translation requires more powerful computing power than statistical machine translation. With the continuous development of artificial intelligence technology, new technologies for natural language processing will change the industrial construction of translation, and the translation industry chain may experience a major industrial change.

2. Text Analysis

Text analysis also known as text mining or text analytics. Although the terms are sometimes used interchangeably, there is a difference between text analysis and text mining. Meanwhile, text analysis provides you with more detailed quantitative data to help you make informed decisions. Text analysis, when applied to unstructured information (also known as open-ended input), provides insight into trends, patterns, and consumer attitudes, allowing companies to identify and prioritize ways to improve customer experience.

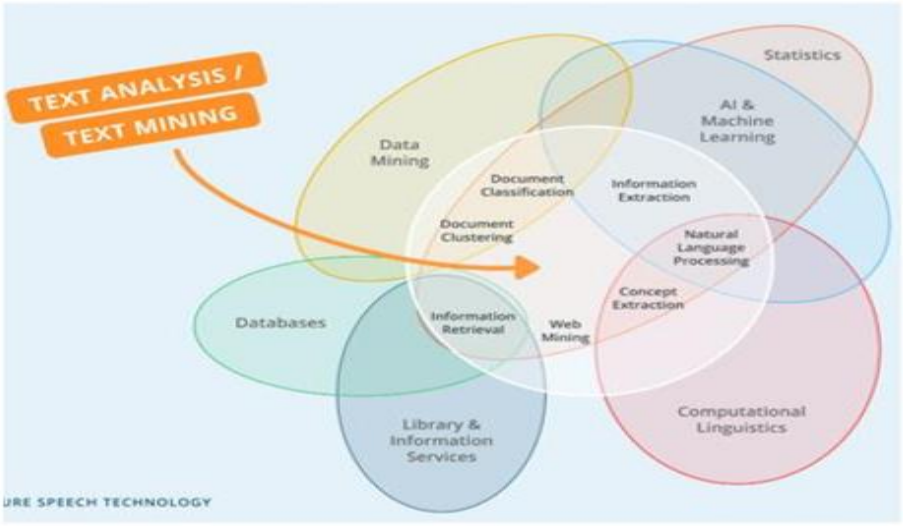


Fig2: Domain and Subdomain of Text Analysis

Natural language processing and text analysis are critical components of artificial intelligence and many other sorts of technology. Text analysis turns unstructured text within documents and databases into normalized, organized data that may be utilized for analysis or to train machine learning algorithms using natural language processing (NLP). The evolution of technology is responsible for assisting individuals in gaining positive results from its use. Text analysis and natural language processing in artificial intelligence are convicted in this study.

During the research, the content analysis approach was used to analyze the literature and studies that were already available in order to have a better understanding of the topic. The study's conclusion was that text analysis and natural language processing are more important in artificial intelligence than previously thought.

Text analysis is used to explore textual material and derive new types of variables from raw data that can then be displayed, filtered, and utilized as inputs for models and other statistical procedures [21]. Natural language processing and text analysis are combined in a variety of applications, including [22].

- Investigation discovery
- Subject matter expertise

Text analysis in artificial intelligence is definitely moving in this direction, breaking down the ideas and structure of the text such that the person running it can readily understand it [23].

LITERATURE REVIEW

Kongthon, 2009 mentioned the use of natural language processing in artificial intelligence to construct an online tax system. This particular implementation was demonstrated in order to persuade people to use natural language processing and text analysis in artificial intelligence in order to secure the future. The majority of high-level natural language processing applications incorporate components that have been imitating intelligent behavior, such as apparent natural language understanding. Following a review of methods in this field by Jean, 2014. He propose a strategy based on significance sampling, which allows us to use a large-scale vocabulary without increasing the NMT model's training complexity, to solve machine translation. Then, using a single neural network configured to improve translation performance, they suggest an approximation training strategy based on biased sampling that will allow you to train an NMT model with a much wider target phrasing.

Text mining is the process of examining a large quantity or collection of textual materials in order to generate new information. Text mining, for example, aims to identify important information in text by transforming it into data that may be used for further research. One of the technologies used in text mining to attain this goal is natural language processing (NLP). Any natural language processing programme must have a thesaurus, lexicon, ontology, and current entities to work. (Fisher, 2010)

Social media monitoring is one of the most common applications of text mining and NLP, in which a pool of user-generated content is examined to determine mood, emotions, and awareness related to a topic [23].
Natural Language Processing (NLP) is a set of techniques for analyzing such material in order to deduce the underlying meaning (NLP). Machine Learning (ML) techniques are used to create apps that classify, extract structure, summarize, and translate data, which can be speech, text, or even a picture [23].

Research Methodology
Content analysis is a way for analyzing existing literature and having a thorough understanding of Text Analytics and Natural Language Processing in Artificial Intelligence. Various research papers and earlier analyses were considered, and literature that was previously available was extracted using keywords. The material was extracted from various literatures, and after that, it was verified for accuracy and reliability using the relevant sources. Details were enclosed in paper after accuracy testing, and irrelevant material was deleted. Following the addition of important information in the article, a later examination of this particular topic was completed.

Conclusion
According to the findings of the study, NLP is a far superior way to other methods since it is capable of identifying both text and speech, whereas text mining solely deals with the evaluation of text quality. Text mining requires knowledge of different aspects such as cosine similarity or feature hashing, text processing, and programming languages such as Perl or Python. NLP systems require less knowledge of skills such as NLTK and neural network proficiency, whereas text mining requires knowledge of different aspects such as cosine similarity or feature hashing, text processing, and programming languages such as Perl or Python. Text mining requires a thorough understanding of statistical methods. According to research conducted over the previous ten years, NLP is being utilized more frequently than other statistical methods since it is more practicable, simple to use, and requires less knowledge.

3. Conversational & Question Answering Systems

Introduction
Conversational systems allow us to have natural-language conversations with automated systems via voice or text interface. They aid in the automation of complicated procedures in an organisation while providing users with round-the-clock support. Chatbot and Virtual Assistants are the most frequent kind of talking gadgets. Today, banks, e-commerce, social media, and other self-service point-of-sale systems use these two devices to give a variety of services to its customers.
In order to react to typical open-ended queries found in a vast collection of instructive texts or data. To accomplish this, we must first comprehend the types of questions that are being posed, as well as possess a basic understanding of the process. Knowledge of semantic and syntactic analysis, as well as crafting that type of question with probable valid replies, will be among the requirements. As a result, this study provides a clear picture of how question answering is done and introduces a machine learning approach. Also, in a different data set of knowledge, ways are a significant way of supplying solutions to closed domain.

Question Answering System

Posing a human-language query and providing a proper response a specific inquiry may be posed at times, although an open-ended question may also be asked. Even more difficult questions have arisen as a result of recent study.

Literature Review
Question answering software is a branch of computer science that deals with natural language processing and information retrieval. Question Answering entails the creation of defined systems that respond to questions posed by humans in a natural language format. A computer software creates its own answers by asking or querying a structured database of information or knowledge, which is commonly referred to as a knowledge base, during the implementation phase. It may even be able to extract some answers from an unstructured dataset of information or knowledge.

There were two QA systems in the beginning, BASEBALL and LUNAR. Knowledge bases were established in the 1970s to stream knowledge areas. The question-answering system creates a connection to these expert systems. Except for the fundamental design, these expert systems are similar to modern QA systems. Computational linguistics emerged in the 1970s and 1980s, resulting in increasingly helpful and effective initiatives in the fields of question answering and text comprehension. EAGLI, a new quality assurance system, was created for health and life benefits.

Question Answering Methods

The higher the collection data size, the better the performance. QA is heavily reliant on a robust search termed corpus, thus the larger the collection data size, the better. The focus of data redundancy in a vast collection is that small bits of information or data are to phrased in different ways as their by having all the right information that appears in many forms and thus it forms too much work on the QA system to do the bigger or complex works that must be understood and also having the correct values nearby then the incorrect values can be corrected and rectified. The QA system is entirely based on reasoning power, and there are a plethora of question-answering systems that have been established and developed in Prologue, a logic programming language tied to artificial intelligence.

Types of QA System:

Closed domain question answering

It is focused on a certain topic or domain. Because NLP systems excel at locating precise topic searches and extracting responses, it may be easier to ask questions and receive responses. Closed domains can easily determine the location of the questions in reference to a specific node, allowing them to retrieve the answers to the questions. Procedural questions will be less successful than descriptive inquiries. Machine reading applications are also being developed in the medical field. Eg: Alzheimer's disease.

Open Domain Question Answering

In the field of information retrieval, it is stated that in an open domain QA system, the purpose is to provide the correct answers to the user's intended questions. As a result, the returned material is distributed in the form of brief fragmented text rather than documents. As a result, the QA system identifies the answers using information retrieval, computational linguistics, and knowledge representation of data to identify the answers to the questions. The QA system uses a natural language query rather than a list of keywords as an input. Thus, having the main input as a natural language question is more simple and user-friendly, but it is more difficult to implement due to the fact that there are numerous question types and it is difficult for the system to select the correct one from among them in order to provide the appropriate correct answer. Introducing each inquiry type to a question is a difficult and time-consuming task. The entire retrieval process is based on determining the correct question type, which in turn leads to the correct answer type, which aids in determining the correct appropriate value.

Conclusion

The machine learning technique to question answering is outlined in this work. This paper depicts question answering in a closed domain, allowing the listener to quickly respond to those queries. As a result, the closed domain question-answering strategy has a high level of user acceptance. In the future, we will endeavor to reduce the amount of knowledge about semantic analysis and its value, as well as to handle some other concerns and difficulties that arose over those days, and we will undoubtedly resolve all of these issues in future renders.

References

[1] Ise ,Orobor Anderson. "Integration and analysis of unstructured data for decision making: Text analytics approach." International Journal of Open Information Technologies 4, no. 10 (2016).

[2] Fisher, Ingrid E., Margaret R. Garnsey, Sunita Goel, and Kinsen Tam. "The role of text analytics and information retrieval in the accounting domain." Journal of Emerging Technologies in Accounting 7, no. 1 (2010): 1-24.

[3] S. Vijayarani1, J. Ilamathi and Nithya, "Preprocessing Techniques for Text Mining - An Overview", International Journal of Computer Science & Communication Networks, Vol.5, issue.1, pp. 7-16 7 ISSN: 2249-5789

[4] N. Kaur1, V. Pushe and R Kaur,"Natural Language Processing Interface for Synonym", International Journal of Computer Science and Mobile Computing, Vol.3 Issue.7, July- 2014, pp. 638-642 ,ISSN 2320–088X

[5] P. Jackson and I. Moulinier, "Natural Language Processing for Online Applications": Cambridge University press, New York.2012, page 7-9.

[6] K. W Church and L.F Rau," Commercial applications of Natural Language Processing". *Communication of the ACM*, vol 38, No. 11, November 1995

[7] E.D. Liddy, Natural Language Processing, 2001.

[8] ACL 2018: 56th Annual Meeting of Association for Computational Linguistics https://acl2018.org

[9] Predictive Analytics Today: www.predictiveanalyticstoday.com [accessed in Dec 2018]

[10] Day, M. Y. (2020). Artificial Intelligence for Text Analytics

[11] Cavazza, Marc, Srikanth Bandi, and Ian Palmer. "Situated AI in video games: integrating NLP, path planning and 3D animation." In AAAI 1999 Spring Symposium on Artificial Intelligence and Computer Games, pp. 6-12. 1999.

[12] S. Jusoh and H.M. Alfawareh, "Natural language interface for online sales", in Proceedings of the International Conference on Intelligent and Advanced System (ICIAS2007),Malaysia: IEEE, November 2007, pp. 224-228

[13] E.K. Ringger, R.C. Moore, E. Charniak, L. Vanderwende, and H Suzuki, "Using the Penn Treebank to Evaluate Non-Treebank Parsers", In Proceedings of the 2004 Language Resources and Evaluation Conference (LREC), 2004, Lisbon, Portugal.

[14]. Predictive Analytics Today: www.predictiveanalyticstoday.com [accessed in Dec 2018]

[15] Le, Quoc V. and Mike Schuster. 2016. Neural Network for Machine Translation, at Production Scale.n.d.Sep.27,2016.https://research.googleblog.com/2016/09/a-neural-network-for-machine.htm.

[16] Hutchins, W. John. 1995. Machine Translation: A Brief History. In E. F. K. Koerner and R. E Asher (eds.), Concise History of the Language Sciences: From the Sumerians to the Cognitivists. Oxford: Pergamon Press.

[17] Weaver, Warren. 1955. Translation. In William N. Locke and Andrew Donald Booth (eds.), Machine Translation of Languages: Fourteen Essays. Cambridge: MIT Press.

[18] Foster I, Kesselman C, Nick J, et al. Computer Grid Services for Distributed System Integration[J]. IEEE Computer, 2002, 35(6)

[19] Bahdanau D, Chok, Bengio Y. Neural Machine Translation by jointly learning to align and translate [J]. arXiv:1409.0473,2014

[20] Junczys-Dowmunt M, Dwojak T, Hoang H. Is Neural Machine Translation Ready for Deployment? A Case Study on 30 Translation Directions [J]. arXiv:1610.01108v2,2016

[21] Ahonen, Helena, Oskari Heinonen, Mika Klemettinen, and A. Inkeri Verkamo. Applying data mining techniques in text analysis. Report C-1997-23, Dept. of Computer Science, University of Helsinki, 1997

[22] Lee, L. (2002, July). A non-programming introduction to computer science via NLP, IR, and AI. In Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics (pp. 33-38).

[23] Moreno, A., & Redondo, T. (2016). Text analytics: the convergence of big data and artificial intelligence. IJIMAI, 3(6), 57-64.

[24] Mnasri, Maali. "Recent advances in conversational NLP: Towards the standardization of Chatbot building." arXiv preprint arXiv:1903.09025 (2019).

[25] Lalwani, T., Bhalotia, S., Pal, A., Rathod, V., & Bisen, S. (2018). Implementation of a Chatbot System using AI and NLP. International Journal of Innovative Research in Computer Science & Technology (IJIRCST) Volume-6, Issue-3.

[26] Hirschman, Lynette, and Robert Gaizauskas. "Natural language question answering: the view from here." natural language engineering 7.4 (2001): 275-300.