

Intelligent Computing with Natural Language Processing

Jagjit Kaur*, Simarjot Kaur, Anu

Department of Computer Science and Engineering, Chandigarh Group of Colleges, College of Engineering, Mohali, Punjab, India

Abstract

Natural language processing is a branch of computer science and artificial intelligence which involve machines or robots to understand and process the language that humans speak or which is concerned with interaction between computer and human languages. Natural Language Processing puts great efforts for making computer interfaces so that it will become easier to operate the machine through natural language rather than commands. There are various languages that support NLP like Python for example. Google assistant works on voice recognition for the natural language processing. A programming language makes our job easier to train a machine. The vast computer science field deals with the interactions between human language and computers that call it as Natural Language Processing, or NLP. NLP is a component of artificial intelligence (AI). NLP is a technique to examine, understand, and evaluate meaning of human language in a productive manner. By using NLP, producers generate structure knowledge to perform specific tasks such as automatic summarization, translation, named entity recognition, sentiment analysis, voice recognition, and topic segmentation. NLP is compatible as a hard problem in computer science. Human language is rarely precise, or easily spoken. To understand human language is to understand not only the words, but the main focus is on pattern transformation or pattern matching. Although languages are considered being the easiest things for humans to learn, but on the same phase, it is considered to be a challenging task for the machine to understand human languages and to process it.

Keywords: Natural Language Processing (NLP), Artificial Intelligence (AI)

***Author for Correspondence** E-mail: jajjitk.coecse@cgc.edu.in

INTRODUCTION

Figure 1 shows phenomenon of natural language processing.

Components of Natural Language Processing

Natural Language Understanding (NLU)

In this, there is ambiguity, means multiple solutions. It consists of:

1. *Lexical (Word Level)*: during natural language understanding, system may read the words of the sentence as noun or may be verb; so, gives multiple solutions. e.g. 'Water' is a noun and a verb also. So what should be perceived by the computer?
2. *Syntactical (Parsing)*: It works on little bit part of the sentence e.g. "call me a cab". So, in this in our terms we want the taxi but the system may take it as now my name is cab. So in this way, it understands multiple solutions.

3. *Referential*: In this, the system may interpret the full sentence and may not be able to relate it with another sentence. So it replies with multiple solutions. e.g. 'John met David. He said I am hungry'. Now "He" is referring to whom? John or David? So, this is a great problem [1, 2].

Natural Language Generation (NLG)

The system understands the natural language and now generates the answer. It consists of:

1. *Text Planning*: For this, system understands each and every word and generates the text.
2. *Sentence Planning*: After understanding the text, system generates the meaningful sentence by arranging the words or with the help of text planning.
3. *Text Realization*: After sentence planning, system generates the structured form of the sentence and finally generates the output (Figure 2).

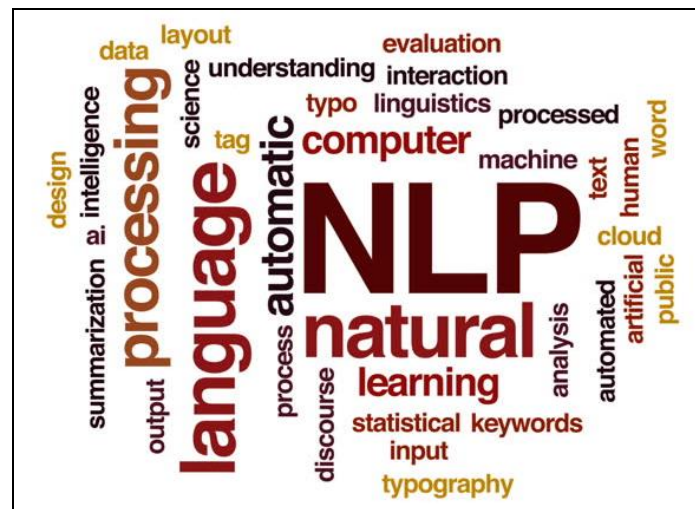


Fig. 1: Phenomenon of NLP.



Fig. 2: Natural Language Generation.

Now we will see the steps involved in successfully processing and analyzing the sentences for right understanding. These 4–5 steps are sequential and execute one after another (Figure 3). The processed texts can be used to analyze the sentiments of the writer, or can be used to create summaries, etc. [3, 4].

APPROACHES OF NATURAL LANGUAGE PROCESSING

There are four approaches of natural language processing:

Distributional

The first approach includes the deep use of machine learning and deep learning. This approach converts content into word vectors for mathematical analysis and performs tasks such as part-of-speech tagging like, is this a noun or a verb? The sentences are parsed and then analyzed, that is, it noun or a verb, called dependency parsing [5].

Frame-based

Basic role of frame based approach is: Sentences that are syntactically different but semantically identical.

Model-Theoretical

The third category of approach includes model-theoretical approach. It is further divided into two concepts: model theory and

compositionality. Model theory is the idea that sentences refer to the word, as in the case with grounded language. In compositionality, meaningful part of a sentence can be combined to conclude the whole meaning.

Interactive Learning

In this category, the user says a sentence step by step and indicates to the computer what the result of the execution should look like.

ALGORITHM DESIGN FOR NATURAL LANGUAGE PROCESSING

For a system to understand human language, a technology must also understand the grammatical rules, meaning and context, but also colloquialisms, slang and acronyms used in a language. Natural Language Processing algorithms are based on the computer understanding of human language and behavior. NLP algorithms are based on statistics and include deep learning and machine learning [6, 7]. Lexicon is put at the core status when semantics technology is used in NLP in order to understand the language. Following algorithms are useful for the natural language processing:

1. Word Alignment in Machine translation: Bitext.

2. Spell Checker: Edit Distance, Soundex.
3. Parsing: CKY algorithm and other chart parsing algorithms.
4. Document Classification: SVM, Naive Bayes.

NLP used in Chat Bots (Flow Diagram)

NLP used in chat bots and the flowchart is given below in Figure 4.

NLP using Programming Languages:

Python

Python is a modern high level language that can be used for the analysis of texts by NLP. It comes with many libraries supporting it. "Text Blob" is one of the libraries. It provides an API for common NLP [8, 9].

Text Blob Features

1. To understand "who" in the sentences, noun phrases extraction can be performed by it. It has an inbuilt function for `it`.
2. We can find out the parts of speech in the sentences.
3. Can pluralize or singularize the words in sentences.
4. Sentiment Analysis: this is the most interesting feature in Text Blob; it can sense the mood of a sentence by rating it from -1 to 1 (known as polarity). And it can be used in movie reviews, analyzing the moods of twitter users by their tweets.
5. Can summarize a big paragraph.
6. Translation of language from one to another.

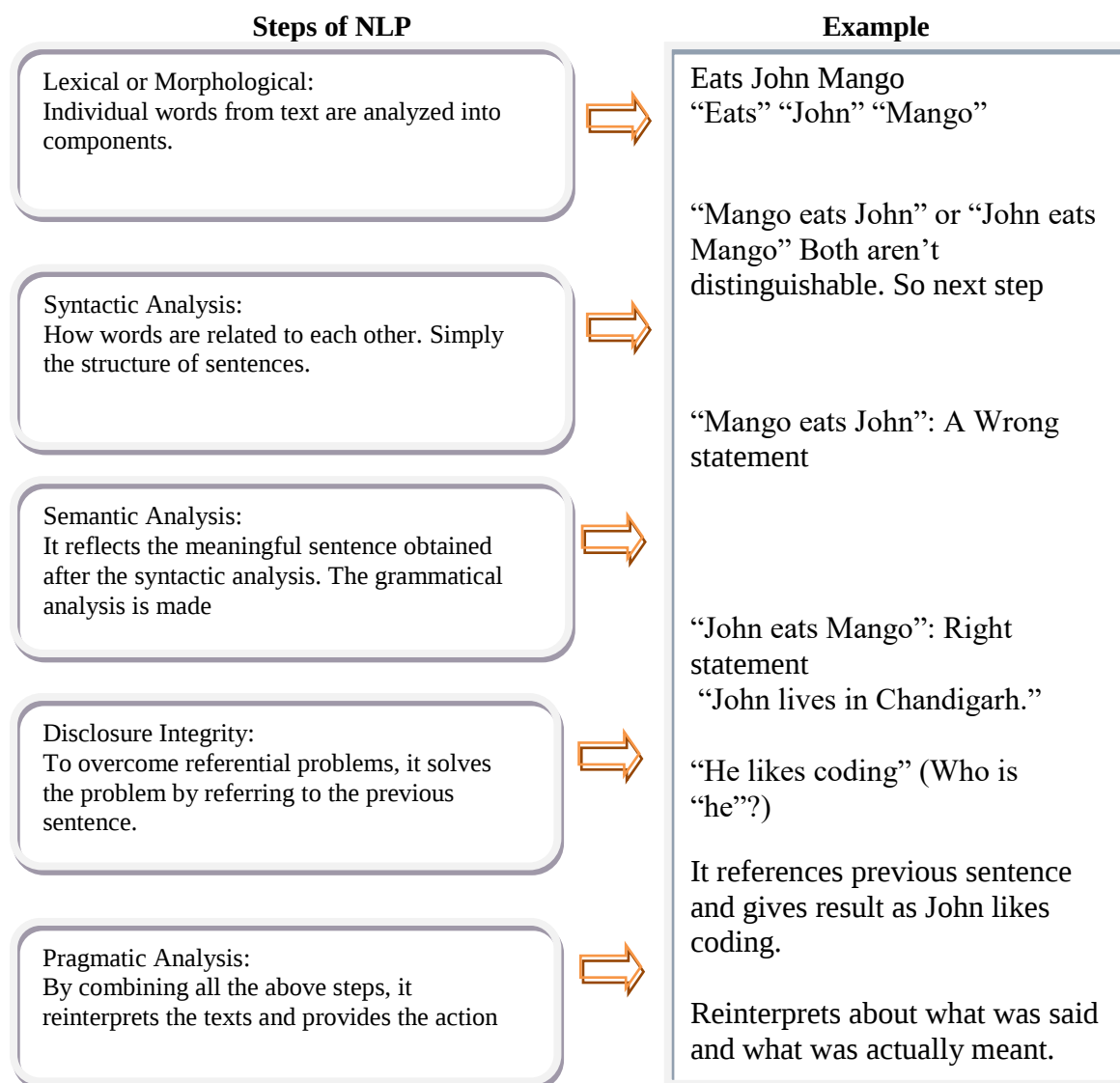


Fig. 3: Natural Language Processing Concepts with Example.

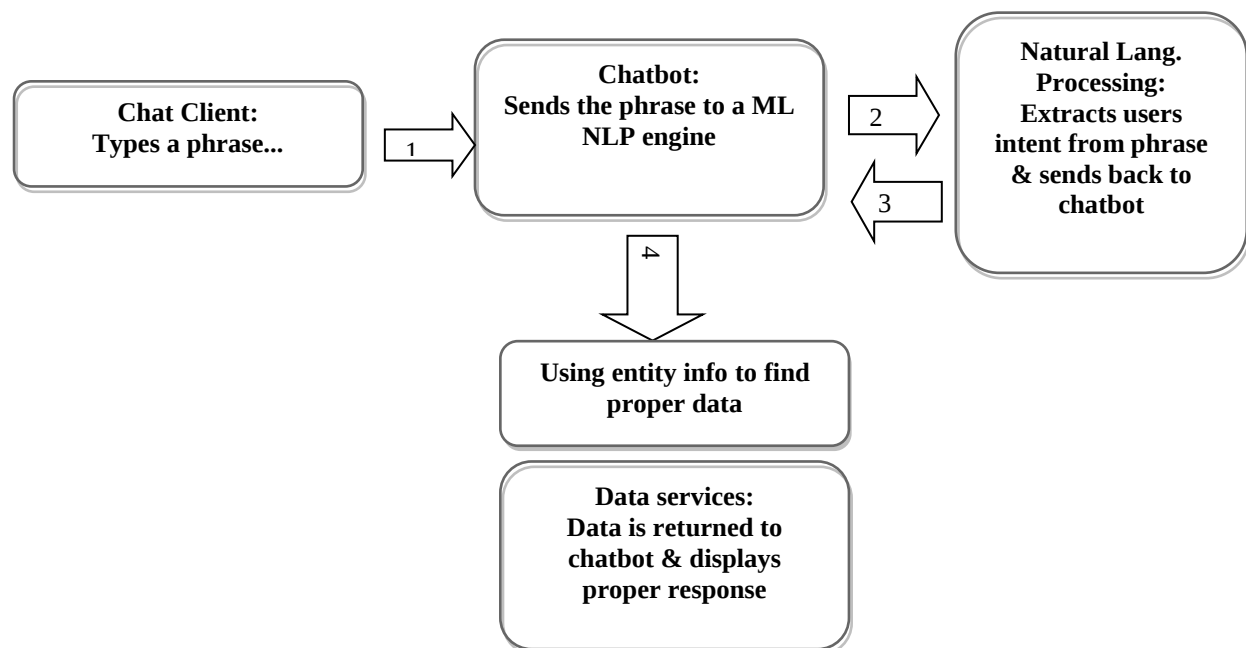


Fig. 4: Natural Language Usage-CHAT BOT.

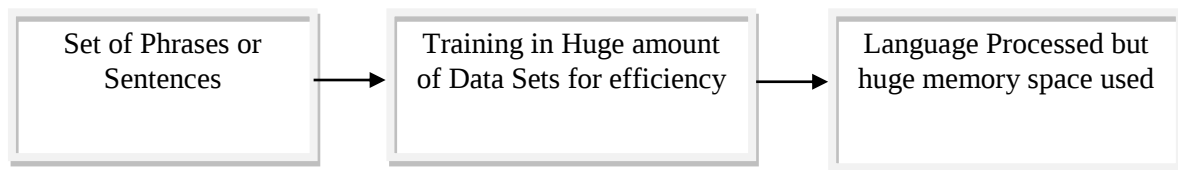


Fig. 5: Natural Language Problems.

APPLICATIONS OF NLP

Automotive Virtual Assistants

NLP works on automotive virtual assistants like BMW. The working is like when the user parks the car; then with the help of software application (cellular phone) and voice recognition just to instruct the AI to send a text message from the user's current location to car. It allows user to access the device and services through voice instruction, like navigation, music, message dictation, calendar, weather, and social media [10].

Health Care Services

Intermountain is one of the methods using a self-developed, Java-based NLP program to identify what kind of illness a patient is suffering from. Intermountain has used NLP with the help of which it identifies stroke patients, cancer patients, heart failure patients etc. There is also the process going for the creation of robots used for the proper medical care of the patients. Healthcare natural language processing is using enable automated

case finding for the best candidates for clinical trials and medical research.

Customer Services

Natural Language Processing deals with the customer services like services for mobile phone. When user has any query related to services, it deals with by understanding the problem of the user. It also gains more and more attention with the raise of Chat Bot.

- a) Understanding customer feedback: Finding patterns of sentiment (positive or negative) by very heavy datasets used to train the machines in huge volumes of customer interactions, and using these analysis, helps in improving business and areas of both, shortcomings, improvement and excellence.
- b) Speech processing: With the help of NLP programs, you can book a hotel, railway tickets or flight tickets on phone.
- c) Data extraction: Find out the names, address of the people from a huge data who are in specific events.

- d) All queries answered: Find answers of questions by NLP through texts in datasets.
- e) Summarizing: Generate short summary from multiple news articles.

Limitations

The greatest challenge to NLP is to main focus on pattern recognition representing a sentence or group of concepts with absolute precision. The configuration of software's present and hardware's limit this challenge. There is also the need of realistic amount of data necessary to perform NLP at the human level which requires a memory space and processing capacity; that is the major problem even with the most powerful computer processors [11, 12].

As the feature of NLP system developed in cars but if the voice of the person is copied then unauthenticated use of his/her voice to unlock the car or change the functionality. So the prior focus of NLP is to work on the pattern in which user speaks which cannot be copied by the unauthenticated user (Figure 5).

CONCLUSION

Like everything has two aspects, Machine learning, artificial intelligence and intelligent computing too have these. In the present world, humans have made it possible to make such machines who can handle the tasks same way as themselves, can show emotions and gestures. There is a YouTube page named Boston Dynamics where such quality robots are made that can perform all human tasks. But training them requires high skills; the machines are to be trained. The AI like Google assistant is the best open source assistant in the android phones but still has some shortcomings that have to be patched for a better output. A more powerful AI named GOOGLE DUPLEX will be soon coming to our devices. It was seen in Google IO that can make some appointments on your behalf by same way a human speaks. That is very powerful and interesting. The pattern recognition has to be improved for better efficiency of many AIs. Large amount of datasets are to be used for highly efficient AI machines. The human language and their

emotions are to be well processed by the NLP techniques, which otherwise can give unwanted results. The data scientists are working upon it to improve the efficiency.

REFERENCES

1. Cambria E, White B. Jumping NLP Curves: A Review of Natural Language Processing Research. *IEEE Comput Intell Mag.* 2014; 9(2): 48–57p.
2. Mikolov T, Karafiat M, Burget L, *et al.* Recurrent Neural Network Based Language Model. In *Interspeech.* 2010; 2: 3p.
3. Mikolov T, Sutskever I, Chen K, *et al.* Distributed Representations of Words and Phrases and Their Compositionality. In *Adv Neural Inf Process Syst.* 2013; 2: 3111–3119p.
4. Socher R, Perelygin A, Wu JY, *et al.* Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).* 2013; 1631: 1642p.
5. Collobert R, Weston J, Bottou L, *et al.* Natural Language Processing (almost) from Scratch. *J Mach Learn Res.* Aug 2011 12: 2493–2537p.
6. Goldberg Y. A Primer on Neural Network Models for Natural Language Processing. *J Artif Intell Res.* 2016; 57: 345–420p.
7. Bengio Y, Ducharme R, Vincent P, *et al.* A Neural Probabilistic Language Model. *J Mach Learn Res.* Feb 2003; 3: 1137–1155p.
8. Mikolov T, Chen K, Corrado G, *et al.* Efficient Estimation of Word Representations in Vector Space. *arXiv preprint arXiv:1301.3781.* 2013.
9. Weston J, Bengio S, Usunier N. Wsabie: Scaling up to Large Vocabulary Image Annotation. In *IJCAI.* 2011; 11: 2764–2770p.
10. Socher R, Lin CC, Manning C, *et al.* Parsing Natural Scenes and Natural Language with Recursive Neural Networks. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11).* 2011; 129–136p.

11. Turney PD, Pantel P. From Frequency to Meaning: Vector Space Models of Semantics. *J Artif Intell Res.* 2010; 37(1): 141–188p.
12. Cambria E, Poria S, Gelbukh A, et al. Sentiment Analysis is a Big Suitcase. *IEEE Intell Syst.* 2017; 32(6): 74–80p.

Cite this Article

Jagjit Kaur, Simarjot Kaur, Anu. Intelligent Computing with Natural Language Processing. *Journal of Artificial Intelligence Research & Advances.* 2019; 6(2): 63–68p.