Madhya Pradesh Journal of Social Sciences
A biannually Journal of M. P. Institute of Social Science Research, Ujjain
ISSN: 0973-855X (Vol 29 No 1, January 2024)

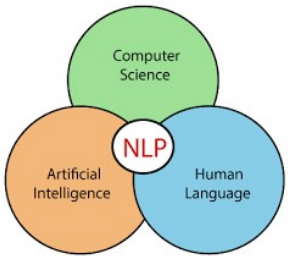
NATURAL LANGUAGE PROCESSING – AN EMERGING FIELD

Dr. Pragati Bhatnagar[1], Dr. Manish Bhatnagar[2]
Assistant Professor, Acharya kalu Kanya Mahavidhyalaya, Jain VishvaBharti Institute Deemed University, LADNU
Assistant Professor, Department of Education,, Jain VishvaBharti Institute Deemed University, LADNU, Rajasthan

Abstract
Natural Language Processing (NLP) means to teach machines to understand and communicate in natural languages. Natural languages are those languages which are spoken commonly by group of persons for communication. Natural languages include Hindi, English, Marathi ,Punjabi etc. There are many application of NLP in our daily life like chatbot, search engines, email-spam filters, alexa etc. NLP is an emerging field of computer science and artificial intelligence and is expected to make revolutionary changes for the benefit of mankind. This paper discusses meaning of NLP, its goal, need and applications in various fields. Further it discusses common NLP tasks, approaches to NLP and the major challenges which are responsible for hindering NLP to grow to its full potential.

Introduction
Natural Language processing is an emerging area in the field of computer science which uses artificial intelligence and has many real life applications like chatbots, search engines, siri, alexa etc. NLP means to teach or train the machine to communicate with human beings in their common speaking language. The work in the field of NLP started near about 50 years back and is still progressive. Initially heuristic approach was used in the field of NLP but with the emergence of machine learning and deep learning this field has started blooming .Still there are many challenges in the field of NLP which are responsible for its slow growth Many NLP engineers and practitioners are working continuously to overcome these challenges and a time will come when NLP will bring revolutionary changes in human life.

Definition of NLP-



Natural language processing (NLP) is a subfield of linguistics, computer science and artificial intelligence concerned with interactions between computers and human language, in particular how to program computer to process and analyse large amounts of natural language data**.**

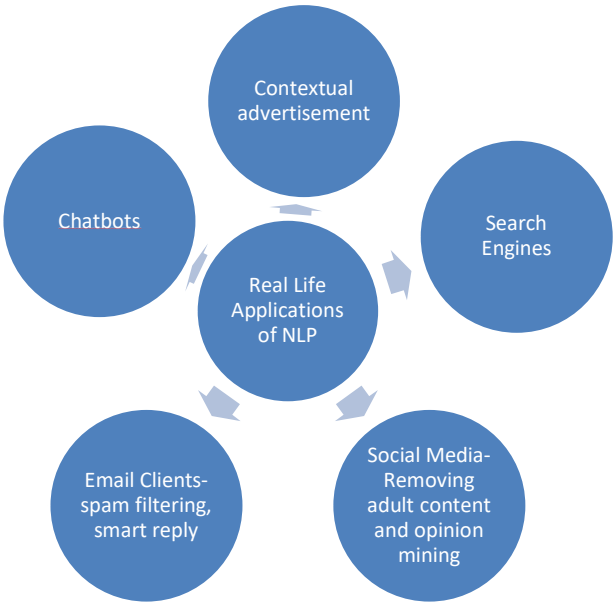
Goal of NLP- The goal of NLP is to teach machines to understand natural language. Natural language processing means machine should be able to communicate with us in natural language. It means

whatever instructions we give to machine will be in natural language and machine should also reply us in natural language.

Need of NLP- To understand the need of NLP we should know what is Natural Language? In neuropsychology, linguistics, and philosophy of language, a natural language or ordinary language is any language that occurs naturally in a human community by a process of use, repetition, and change without conscious planning or premeditation. Natural Language can take different forms such as speech or signing. Natural languages are distinguished from constructed and formal languages such as those used to program computers or to study logic. So natural languages are the languages in which we communicate. It may be Hindi, English or any other regional language.

If we compare the growth of human beings to other animals we find that human beings have progressed very much as compared to other species. If we see evolution of human being we find that our ancestor were similar to monkeys but slowly and slowly we progressed and now a days we are going on moon and making big technological developments. The first reason behind this progress is our communication. Using natural languages human beings share their ideas with others; they pass on knowledge to their next generation in the form of speeches, written text, books, literature etc. The second reason behind the progress of human being is industrial revolution. Man made many machines which made the life easy. Many transportation vehicles, electricity and many other inventions transformed the life of human being. Then came the era of computers and internet which made revolutionary changes in human life. Now there is need of one more revolution which will totally change the life of human beings and that is the era of Natural Language Processing where human beings and machine will communicate and respond to each other in a similar way as human beings do with one another.

Real life Applications of NLP- There are many real life applications of NLP. Common are Google Assistant, Siri, Cortona etc. Other major real life applications of NLP are

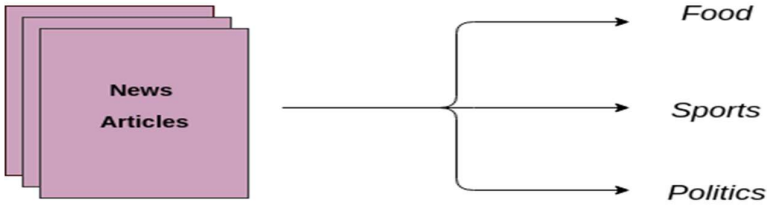


Contextual Advertisement- Previously 20-25 years back when we used to watch matches and other programmes on TV , all the viewers of the country were shown same advertisements. Company used to assume that show the same advertisement to all and those who want to buy the product will purchase it. Now we have NLP and we can process how persons behave, how is their personality, we can show

targeted advertisement to persons. You must have noticed that depending on your instagram profile or facebook profile you are shown different advertisement then your friend. Company used to analyse your profile, posts and comments and then decides whether you are interested in field of sports, education,cosmetics, etc. Depending on your field of your interest targeted advertisement are shown to you. If sometimes we chat on whatsapp with someone on a topic and the switch to some other application we are shown advertisement on the same topic on which we were chatting.

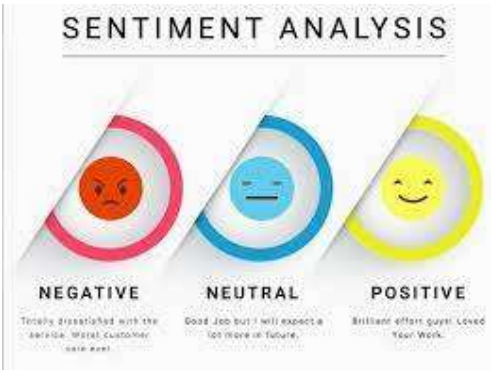
Common NLP Tasks- There are some common NLP tasks using which you can make NLP applications. They are

- **Text Document Classification**



If there is a text and it needs to be assigned a category this can be done using text document classification. For example you have a full text of a news and you want to assign it a category like food, sports, entertainment, politics etc. You can easily do it using text document classification.

- **Sentiment Analysis**

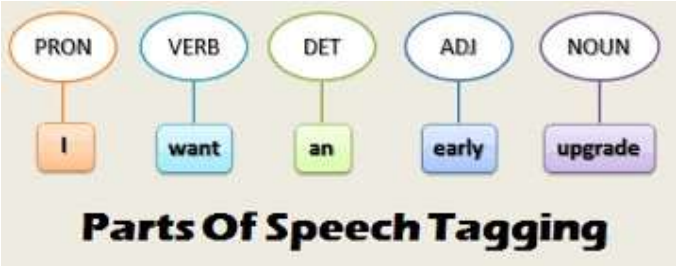


Using sentiment analysis in Social Media or E-commerce you read reviews of people and extract emotions of a people and analyze what they feel about a particular movie or product. Companies use this feature on very large scale. Suppose there is a product of a company and there are one crore reviews of that product. It becomes very difficult to read those review and analyse them. So this can be done by making a sentiment analysis app in NLP such that you can feed all the reviews in the app and it will tell you how much percent of people are talking positively about your product. It will save company's money and time

- **Information Retrieval**

Suppose we have a text and we want to extract entities from it like, time, date, person's name, event name, product name etc. We can do it using information retrieval. Search engine uses informational retrieval internally to search information on web. HP0

- **Parts of Speech tagging**

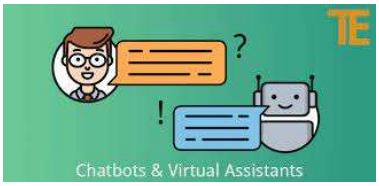


It is a very important text pre-processing step. Here you assign every word of the text a part of speech tag such as noun, pronoun, adjective, verb and determiners etc. It is used if you prepare a question answering system or chat bot where the meaning of a sentence should be clear word by word.

- **Language detection and machine translation-**

It is a very powerful application and is used in google translate where you can speak or type in any one language and it is translated into other language. Here first language is detected on the basis of input given by user and then it is translated on the basis of output language required by the user.

- **Conversational Agents**

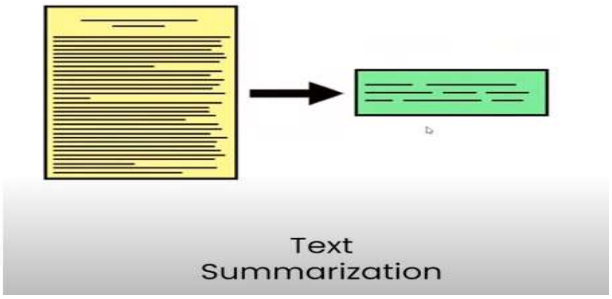


These are important applications of NLP used in chatbots. Basically two types of chatbots are there text based and speech based. Chat bots used in swiggy like applications are text based whereas alexa, siri are examples of speech based chatbots.

- **Knowledge graph and QA Systems-**

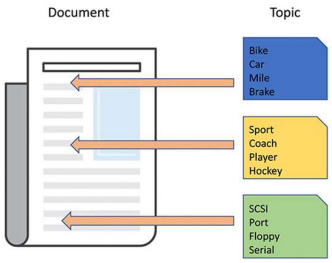
It is one of the most powerful application of NLP. If there is a very large database of thing then its entities can be connected logically and a big knowledge graph can be prepared and after that it can be translated into a question answering system. This application is used by google. The data base of google is very big probably one of the largest in the world. Google has extracted different entities and connected them logically to prepare a big knowledge graph. If any question is asked to google like who is the prime minister of china? It will quickly give the answer using its knowledge graph.

- **Text summarization-**



It is also an application of NLP where a summary of big text is presented few words or summarized form. There is a news application inshot which provides us summary of each news in 60 words using text summarization.

- **Topic Modeling –**

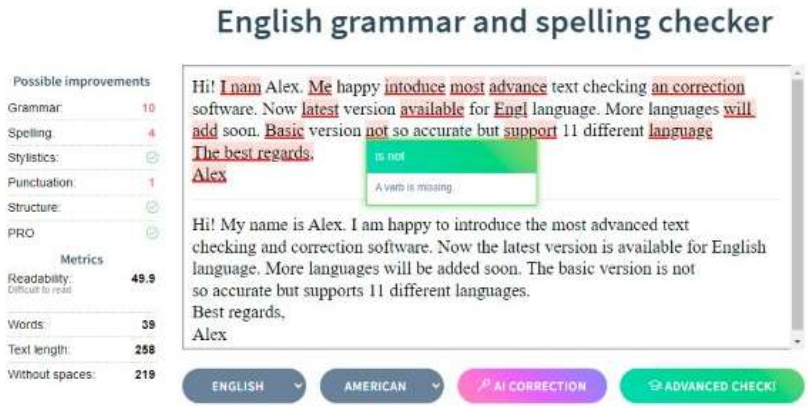


It is also an interesting application of NLP. If you have a very big text you can extract information regarding abstracts topic in the text. Suppose you have a big text about cricket. Now you need to find out it is about what topic on cricket. Is it about IPL, or is it about Sachin, or is it about World cup. This can be done by using topic modeling.

- **Text Generation-**

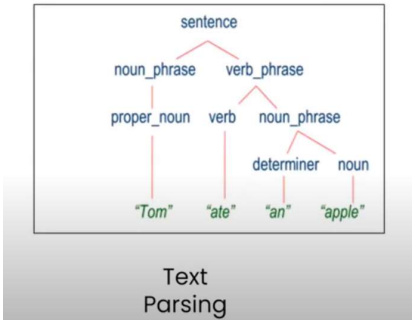
It is very useful feature used by all of us while typing. The keyboards, which are used by us especially when we are using mobiles, predicts on the basis of our previous typing behavior which word is to be typed next. It enhances our speed while working.

- **Spell checking and grammar correction-**



The tools like grammarly are useful tools which identifies our typo and grammatical errors in text, highlights them and give us suggestions to quickly correct our mistakes.

- **Text Parsing-**

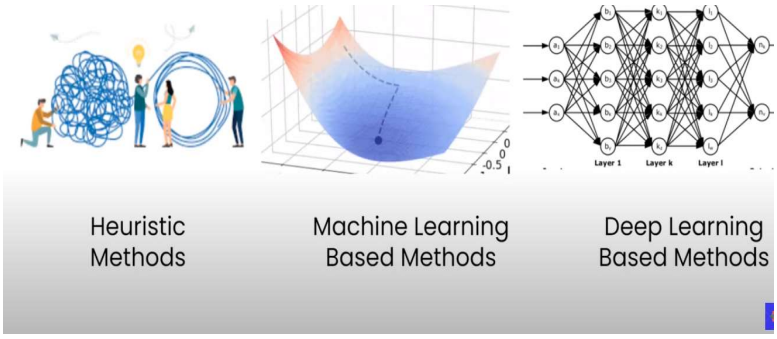


It helps us to grammatically break down our sentence so that we can make machine understand the meaning of the sentence. It is used in making applications like chatbots etc.

- **Speech to text-**

This feature is useful in making conversational agents like Siri and Alexa. Here you speak something which is converted into text and agents try to understand that. Then the reverse process is done and the output or the result is converted into speech. This is also a very important application when you try to talk to a machine and don't want to type.

Approaches to NLP- Different techniques are used to make NLP applications. Basically three methods are used in NLP. These are heuristic approach, machine learning and deep learning techniques.



- **Heuristic Approach**-A heuristic technique is any approach to problem solving or self-discovery that employs a practical method that is not guaranteed to be optimal, perfect, or rational, but is nevertheless sufficient for reaching an immediate, short-term goal or approximation. It is rule-based approach. For example you want to make a sentiment analysis app. A simple rule based approach can be to count number of positive and negative words in the text to be analysed. Popular examples include- Regular Expression, Wordnet and Open Mind Common Sense. Its advantage is that it's very quick and accurate. In this method human beings are responsible for creating rules so heuristic approach provides accurate answers. These approaches came into existence in year 1950 and are used presently also.
- **Machine Learning Approach**- . From 1990's the era of machine learning based method changed the scenario. Machine learning method gave a big advantage over heuristic approach. In heuristic approach rule based method is used. As in many open ended problems we cannot create rules so the role of machine learning approaches came into existence. The rules in machine learning approaches are created by algorithms internally. Relationship between input and output is created on the basis of nature of data instead of creating a general rule. So machine learning approaches are used in wider variety of problems and are more powerful as compared to heuristic approach. Machine learning algorithms learn from experiences
- **Deep Learning Based Methods**- After machine learning came the era of deep learning. Deep learning utilizes artificial neural network. Machine learning methods depends on mathematical model whereas deep learning methods are designed to learn from vast amount of data as they are exposed to very large amount of data. These algorithm identify pattern and relationships in data and use these to produce more

accurate result. In NLP deep learning algorithms learn from relationships between word and phrases and produce highly accurate result.

Challenges in NLP- As natural language processing applications are based on natural languages and as we all know that there are many vague things in natural languages which are understood by human beings but are very difficult to be understood by machine. So there are many challenges to NLP. The main challenges are as below

Major Challenges to NLP

Ambiguity- Sometimes what we speak conveys more than one meaning which could be understood by us but it is difficult for machine to understand. For example a sentence like "I have never tasted a cake quite like that one before" The meaning of this sentence can be predicted by human being on the basis of tone of sentence being spoken by a person. Human being can easily understand whether it tastes good or bad. The same sentence if is to be understood by a machine will be difficult to be interpreted as machine understands the meaning of sentence by word to word translation.
Contextual Word- Sometimes the same word conveys different meaning in different context. For example "I ran to store because I ran out of milk" for human being it is very easy to understand the meaning. It means that as there was no milk in my home I went to store to buy the milk. Here the word ran is used at two different places and conveys different meanings which is difficult to be interpreted by machine.
Colloquialism & Slang-There is some common knowledge hidden in our daily routine talks which humans know. I may talk to a person something which means something different. The person communicating with me will understand the meaning but if it is to be understood by a machine it will not be interpreted correctly. For example "Piece of Cake" and "Pulling your Leg". If a person says "This task is a piece of Cake" it means this is a very simple task and "pulling your leg" mean making fun of somebody which could be well understood by a person but it is difficult for machine to understand as the literal meaning of sentences is different.
Synonym- It means different words conveying same meaning. We use these to express our text more effectively but it becomes difficult for machines to understand.
Irony, Sarcasm and Tonal Difference- Sometimes ironical and sarcastic sentences and our tone of speaking changes the whole meaning of a sentence. This is again understood by human beings but not by machines as they interpret meaning by word to word translation.
Spelling Errors- When we read a paragraph and there are some small spelling errors we are able to understand its meaning in context of paragraph but this is not possible for machine as wrong words do not convey any meaning

Creativity-In texts like poems, dialogues and script sometimes the literal meaning of sentences written is different and it wants to express some different feelings .There is lot of creativity involved in writing such text but for a machine interpretation of meaning based on expressions or feeling is challenging task.
Diversity-This is one of the major challenge for NLP. There are near about 5 to 7 thousand languages in world. To understand grammar of all these languages and to decode them is very difficult. NLP applications have been made for standard languages but if NLP application is to be made for a rare kind of language you will not get enough data to train the machine.

As there are number of challenges in front of NLP only 5% of its potential has been utilized.. NLP practitioners are trying best at their level to utilize potential of NLP to its maximum. In the coming scenario NLP applications will bring a revolution for mankind.

Conclusion

NLP is an emerging area in the field of computer science which uses artificial intelligence and has many real life applications. NLP will enable machines to talk to us i.e., human beings in our common language. Chatbots, search engines, robots are some major applications of NLP. The work in this field started about 50 years ago and is still continuing. Advance developments like machine learning and deep learning are further flourishing this field. Though there are some major challenges which are hindering the growth of NLP but NLP engineers and practitioners are working at large level for developing NLP applications, results of which will be beneficial for entire global world.

References :

Singh N (2021 November 25) Introduction to NLP NLP lecture 1 (Video file). Youtube.https://www.youtube.com/watch?v=zlUpTlaxAKI&list=PLKnIA16_RmvZo7fp5kkIth6nRTeQQsjfX

Pais S. et all (2022), NLP-based platform as a service: a brief review, Journal of Big Data, Springer Open 9:54 https://doi.org/10.1186/s40537-022-00603-5
https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00603-5

https://en.m.wikipedia.org/wiki/Heuristic#:~:text=A%20heuristic%20(%2Fhj%CA%8A,for%20reaching%20an%20immediate%2C%20short%2D
https://en.m.wikipedia.org/wiki/Natural_language
https://en.m.wikipedia.org/wiki/Natural_language_processing