# Approaches of Natural Language Processing & its Applications

**Aniket Vartak[1] Tejas Shinde[2] Vaibhav Talekar[3]**
[1,2,3]Student
[1,2,3]Department of Computer Engineering
[1,2,3]MGMCET, Navi Mumbai. Maharashtra, India

*Abstract*— Form the last decade the usage of web application, text applications and chat bots has been increased. Also demand for the Natural language Processing and text mining is increasing. The natural language processing is effective approach for bringing improvement is such application. In simple words the natural language processing is the ability of computers to analyse the natural language and understand the exact meaning of it. It is related to the human – machine interaction. The NLP is the motivation for the society of those peoples who are not much aware of computer languages but need the computer in their daily use for their survival. [1]The natural language processing comes under the domain of artificial intelligence with the goal of understanding and creating meaningful expressions in human language.it is widely used in the large number of educational purpose applications also like science, linguistics, e-learning, assessment procedures. This paper aims to address the process and the approaches of the natural language processing. The study also highlights the NLP can be utilized in the speech recognition application and for the simple problem formation using LPP.

***Key words:*** Natural Language Processing, Education Applications, Data Mining, Text Mining, Artificial Intelligence, Speech Recognition, Linear Programming Problem

## I. INTRODUCTION

The natural language processing is defined as the automatic manipulation of natural language like speech or text by the software. The natural language can be a text, non – linguistic, human language or speech. The natural language processing is the make the computer able to analyze, understand and synthesize the natural language. It uses artificial intelligence concept to process the natural language. The major applications of NLP are used in Education. The natural language processing follows the approach of natural process of language acquisition integrated with the scientific approach programs. The natural language generation is part of NLP also introduce big enhancement in the human machine interaction.

1) Aims and Objectives

The major aims and objectives of this study is

- To understand the natural language processing.
- To understand the approaches of natural language processing.
- To study the applications of NLP in education.

I Approaches of natural language processing

### *A. Outline of the Study*

This study is based on the natural language processing and its approaches and the applications. In the first section of this report we will study the processing abstract of the natural language processing. It will explain the states in the natural language processing i.e. How this process has been done and in second part the applications and the use of natural language processing in the educational application is introduced.

### *B. Terminologies of NLP*

There are some terminologies used in natural language processing:

1) Natural language – human language, text or speech.
2) Syntax – sequencing the words in meaningful manner.
3) Semantics – analyses the meaning of the words or sentences.

There are 5 phases in natural language processing

*1) Lexical Analysis*

In this phase of NLP the words are separated. The words are analyzed into their components and non-linguistics. It identifies and analyze the structure of the words. The lexical analysis divides the whole data into groups of paragraphs, sentences and words.

*2) Syntax Analysis*

The syntax analysis colorizes the grouped words into their respective types, so that it will be easy for the system to understand the purpose of the word used in the sentence. Some general categories are nouns, verbs, adjectives, adverbs, action, relation, numbers, values etc.

*3) Semantic Analysis*

It finds the exact meaning from the extracted words separated from the syntax analysis. It makes the right sequence of words so that it will be meaningful sequence that make sentence. The grammar rules are also analysis in semantic analysis.

*4) Pragmatics*

It deals with the understanding of the sentence. It checks different situations and conditions and finds the interpretation of the sentence. It affects the meaning of the sentence and the meaning with interpretation affects what action should be performed. E.g. – come here. Should have been request rather than order.

*5) Disclosure Integration*

The meaning of the sentence can be depending on previous or the next sentence and it may influence the meaning of sentence that follow it. This phase checks the previous data and detects the actual meaning of sentence.

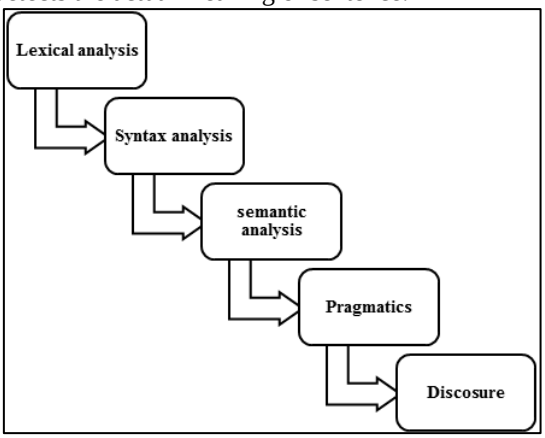


Fig. 1: Framework of the Natural Language Processing

 

The natural language processing is branch of artificial intelligence, which focus on the improvement and the development in the learning process. NLP is can be applied in the education, automated processing AI agents, chat bots and human-machine interaction. The natural language processing is also an effective approach for developing an efficient system of managing linguistic input in the natural settings through various words, sentences and texts. Natural language processing is a widely recognized area in the language learning all over the world. It is implemented successfully in the many languages like Chinese, Japanese, and French. But mostly is uses English due to simplicity of the language and also English is world's most used language.

### *C. Text Mining*

Text mining and data mining is important in the natural language processing. The NLP system once get new data, it stores that data in its database so that it can be used again when similar input is found. This reduces the processing time utilization. Data mining is the process of analyzing data from different perspectives and summarizing it into useful information [3]. It is the process to uncover the hidden information. Text mining, one of the techniques of data mining, is an analysis of data contained in natural language text [4]. Text mining is the processing of constructing information from instructed data. It is also known as intelligent text analysis. Text mining is used to convert the unstructured data into structured data or meaningful information. The data that resides in a fixed field within a record or file is called structured data. This includes data contained in relational databases and spreadsheets, whereas the data that refers to the information that doesn't reside in a traditional row-column database is called unstructured data. Text Mining uses Natural Language Processing to increase the efficiency of mining.

### *D. Approaches used in NLP*

Percy Liang, a Stanford CS professor and NLP expert, breaks down the various approaches to NLP into four distinct categories:

#### *1) Distributional*

This approach includes the large scale statically tactics of machine learning. It turns the content into word vectors for mathematical analysis and perform quite better at part of speech tagging, dependency parsing and semantic relatedness. These tasks don't rely on understanding the meaning of words, but rather on the relationship between words themselves. These systems are flexible, bread and scalable and there can be applied to different types of text without the need for expert encoded domain knowledge. It compares the words to other words or words to sentences or the sentences to sentences and give results in different outcomes. Although distributional methods achieve breadth, they cannot handle depth. Complex and nuanced questions that rely linguistic sophistication and contextual world knowledge have yet to be answered satisfactorily.

#### *2) Frame Based*

A frame is a data-structure for representing a stereotyped situation, explains Marvin Minsky in his seminal 1974 paper called "A Framework For Representing Knowledge." Think of frames as a canonical representation for which specifics can be interchanged.

Liang provides the example of a commercial transaction as a frame. In such situations, you typically have a seller, a buyer, goods being exchanged, and an exchange price.

- E.g. Input – Jack sold the car to tony for $500.
- Output – Seller: Jack; Predicate: Sold; goods: The Car; buyer: Tony; price: $500

Sentences that are syntactically different but semantically identical – such as "Jack sold tony the car for $500" and "tony bought the car for $500 from Jack" – can be fit into the same frame. Parsing then entails first identifying the frame being used, then populating the specific frame parameters – i.e. Jack, Tony, $500.

The obvious downside of frames is that they require supervision. In some domains, an expert must create them, which limits the scope of frame-based approaches. Frames are also necessarily incomplete. Sentences such as "Jack visited the car shop yesterday" and "Jack bought the cheapest car" cannot be adequately analyzed with the frame we defined above

#### *3) Model-Theoretical Approach*

The third category of semantic analysis falls under the model-theoretical approach. To understand this approach, we'll introduce two important linguistic concepts: "model theory" and "compositionality".

Model theory refers to the idea that sentences refer to the world, as in the case with grounded language (i.e. The block is blue). In compositionality, meanings of the parts of a sentence can be combined to deduce the whole meaning.

Liang compares this approach to turning language into computer programs. To determine the answer to the query "what is the largest city in Europe by population", you first have to identify the concepts of "city" and "Europe" and funnel down your search space to cities contained in Europe. Then you would need to sort the population numbers for each city you've shortlisted so far and return the maximum of this value.

#### *4) Interactive Learning*

Paul Grice, a British philosopher of language, described language as a cooperative game between speaker and listener. Liang is inclined to agree. He believes that a viable approach to tackling both breadth and depth in language learning is to employ dynamic, interactive environments where humans teach computers gradually. In such approaches, the pragmatic needs of language inform the development.

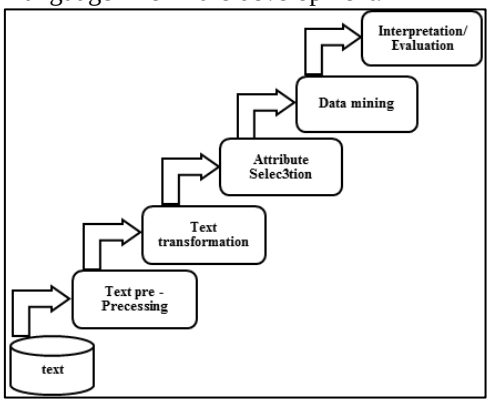


Fig. 2: Text Mining Process

 

## II. EDUCATIONAL & OTHER APPLICATIONS OF NLP

### *A. Teaching about Language*

One of the oldest and still active educational application areas for NLP involves language learning. Basically, the language assessment involves evaluating student's skills in reading, writing or speaking. Use of NLP in language assessment helps in error detecting and correcting in writing such as incorrect preposition usages for deaf students. The standard proofreading tools do not focus on errors that are particularly important for language learners, so for grammatical error detection Natural Language Processing is needed.

### *B. Teaching using language*

The language can also be used as a teaching method. If we consider the traditional method of tutoring it has been found that the students having one on one tutor scores good and more marks then the students under traditional tutoring method. [2]One major difference between human tutors and current computer tutors is that only human tutors participate in unconstrained natural language dialogue with students, which has led to the conjecture that human tutoring might be so effective because of its use of dialogue. In recent years dialogue-based intelligent tutoring systems have thus become more prevalent as one method of attempting to close the performance gap between human and computer tutors.

### *C. NLP in Healthcare Service*

The natural language processing is used for the medical services. In some of the regions of the world there are very less doctors are there. Not everyone can meet physician or sometimes there is no tie to meet. Here the NLP based AI agent system plays important role in the healthcare. Nuance provides various NLP solutions for the healthcare domain, including computer-assisted physician documentation (CAPD) and clinical document improvement (CDI) solutions. Physician documentation is part of medical records that contain patient clinical status, such as improvements or declines in patient health. CDI is the process of improving such healthcare records to ensure improved patient outcomes, data quality and accurate reimbursement. The physicians also agreed with the CAPD clarifications, and updated their patient's documentation accordingly.

### *D. D. NLP for Customer Care Services*

Now a days there is tremendous growth has been noticed in the various businesses providing services. They also provide the help service called customer care service which solves the minor issues related the product or service provided by the company. Previously there were call centers sections kept in the organization to attempt such issues but now there is no need of one to one human serviceman for every issue. The NLP based AI system analyze the speech by the customer and give him appropriate reply. This also reduce the waiting time of the customers till line is connected to the service man.

## III. CONCLUSION

Natural language processing is a branch of artificial intelligence & computer science and it uses text mining to make the interaction between human and computer, though its purpose is to have interaction among natural language of human beings and computers. The language is the text or speech. The natural language processing applications provide a perfect solution to the various problem and difficulties in the educational system which affect the progress of the students and also provide large enhancement in other application which need human – machine or human – human interaction The use of grammar, syntax, and sentence composition can be efficiently utilized through linguistics software systems such as grammar checkers, which are saves times and provides assistance for both teachers and learners. Therefore, there is need for developing effective approach for the social and cultural perspectives. Implementation of NLP is also effective for using the e-learning approach in order to understand and learn from the data available from the electronic sources. There are also future implementations of this research, which can assist in identifying the complex pattern in language. Further research can be conducted to identify its impact in individual learning, understanding of context, and effectiveness of NLP in writing and assessment procedure.

## REFERENCES

[1] Tembhekar Samta, Kanojiya Monika, "A Survey Paper on Approaches of Natural Language Processing (NLP)", Volume III, Issue III International Journal of Advance Research, Ideas and Innovations in Technology 2017
[2] Diane Litman, "Natural Language Processing for Enhancing Teaching and Learning", Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)
[3] Michael Goebel, Le Gruenwald, "A Survey of Data Mining and Knowledge Discovery Software Tools". SIGKDD Explorations. Copyright 1999 ACM SIGKDD, June 1999.
[4] Anuja Bharate, Devendra Gadekar, - "SURVEY PAPER ON NATURAL LANGUAGE PROCESSING". International Journal of Computer Engineering and Applications, Volume VIII, Issue III, Part I, December 14.

 