Review of Applications of Natural Language Processing

Shanet Mathew[1], Shilpa CS[2], Dr. Bijimol T K[3]
[1]Pursuing master's degree program in Santhigiri College of Computer Science, Vazhithala, Kerala, India
Email: *mca2022_shanetmathew@santhigiricollege.com*

[2]Pursuing master's degree program in Santhigiri College of Computer Science, Vazhithala, Kerala, India
Email: *mca2022_shilpacs@santhigiricollege.com*

[3]Assistant Professor in Santhigiri College of computer Science, Vazhithala, Kerala, India
Email: bijitk@santhigiricollege.com

Abstract: *Language is a way of communication and natural languages are those languages that are spoken by the people. Natural Language Processing, NLP, is a branch of artificial intelligence that deals with the interaction between computers and humans using the natural language. The goal of NLP is to read, figure out, understand, and make sense of the human languages in a manner that is valuable. The different ways of a typical interaction between humans and machines using NLP are: a human talks to the machine, the machine (records on a camera or computer) the sound, sound to text (changing from one form, state, or state of mind to another) happens, processing of the text's data, data to sound (changing from one form, state, or state of mind to another) happens and the machine responds to the human by playing the sound file.NLP uses sets of computer instructions to identify and extract the natural language such that the (without rules, schedules, etc.) language data is converted into a form that computers can understand. When the text has been given, the computer will use sets of computer instructions to extract meaning connected with every legal punishment/time spent punished and collect the very important data from them. NLP provides both explanation (of why something works or happens the way it does) and putting into uses for a range of computer programs. In fact, any computer program that uses text is a candidate for NLP. With NLP, computers can guess (based on what's known) and carefully study human language with more meaning than simply digesting and outputting programmed responses.NLP makes greater use and contact with humans in this context and results in more normal conversations.*

Keywords: communication,understanding,generation

1. Introduction

A natural language additionally referred to as everyday language is spoken or written with the aid of people for general purpose communication. A subfield of Artificial Intelligence, Natural Language Processing (NLP) is reserved for computer systems to understand sentences or phrases written in human languages. NLP caters the ones customers who do not have sufficient time to examine new languages or get perfection in it. Languages can be any like Hindi, French, English, and Chinese etc. A language can be defined as a fixed of policies or set of image. Symbol are mixed and used for bringing throughout data or broadcasting the records.

Natural Language Processing is classified into two parts i.e. Natural Language Understanding and Natural Language Generation which evolves the task to recognize and generate the text .Fig 1.1).
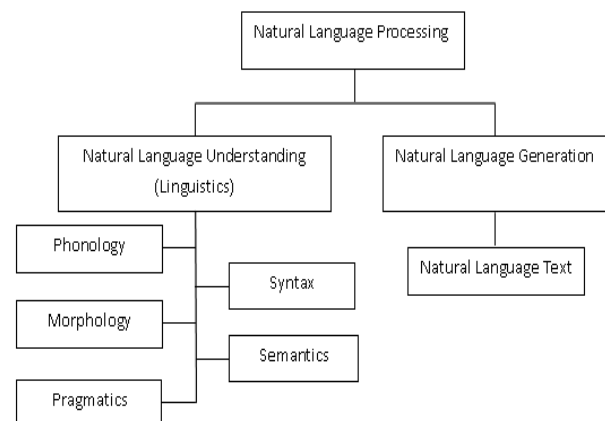
Fig 1.1 Broad Classification of NLP

Linguistics is that the science of language which incorporates Phonology that refers to sound, Morphology phrase formation, Syntax sentence structure, Semantics syntax and Pragmatics which refers to expertise.

2. Applications of NLP

Natural language processing provides both theory and implementations for a variety of applications.Inreality,any

utility that utilizes text is a candidate for NLP. Computer systems can infer and interpret human language with greater significance with natural language processing (NLP) than simply digesting and outputting programmed responses.On this feel, NLP gives greater use and interplay with people and effects in greater natural conversations.NLP makes greater use and contact with humans in this context and results in more normal conversations.

The main areas:
•Massive management of textual information sources:
• For human use
• For automated linguistic tools selection
•Person / Machine interaction.

The following are the most frequent applications that use NLP:

2.1 Machine Translation

Machine Translation is the method of laptop translation of a herbal supply language into some other natural target language. It is a herbal language processing department and has a close affiliation with computational linguistics and interpretation of herbal language.

Neeha Ashraf et al.[1] makes use of English-Hindi MT frameworks reachable on the web for evaluating rule primarily based computing device translation and statistical laptop translation. They are, RBMT system: Anusaarka (http://anusaaraka.iiit.ac.in) and SMT system: Google (http://translate.google.com). The results show that Anusaaraka using RMT approach did pretty properly towards Google with SMT. With regard to syntax and semantics, Anusaaraka correctly recreated Hindi phrases with the RMT approach, preserving the context, grammar and that means of the source English text intact.

According to Li peng [2] all the present distinctive approaches of MT have some disadvantages like grain size is too large for rule based totally method and sparse data for corpus-based approach. Some tips are put ahead after evaluating the advantages and drawbacks of each computing device translation process.focusing on the integration of a range of methods in laptop translation strengthening the role of the semantic analysis in translation introducing linguistic knowledge constructing the of language resources strengthening find out about on evaluation methods

2.2 Question Answering

Question-answering offers the person with both simply the answer text itself or passages that comprise answers. The QA framework receives a question expressed in NL and tries to provide the right answer, now not a textual content containing the answer (usually a fact). For each processing the query and looking for the answer, QA structures need to use NLP techniques.

According to A.Clementeena et al . [3] there are two types of QA system : Closed domain QA and Open Domain QA.
Closed Domain QA: Deals with unique domain or topic. It work out very fast the place the questions are on to a specific node so that it can retrieve the answers for the query that is being asked.
Open Domain QA: In an open area QA gadget aim is to giving out the desirable solutions for the meant questions requested by using the user. Thus the back text is given out in the structure of brief fragmented textual content and not in the shape of documents. Thus the QA machine finds the solutions out of the assist with the information retrieval, computational linguistics and information representation of the information for discovering those answers for the questions given.

O.Ferretet al. [4]have accomplished an evaluation for the choice of the first-rate search engine for QA using Indexal, Zprise and ATT by comparing the results of the three engines for the 200 best files they return And according to the three parameters maintained, the ATT search engine bought the best results: the maximum number of questions for which at least one right report has been returned, the minimum number of questions for which at least one right report has been returned, and the most variety of documents recovered, all of which are blended into questions.

2.3 Information Retrieval

Information retrieval (IR) can be described as a software application that handles the organization, storage, retrieval and comparison of file repository information, particularly textual information.

David D. Lewis et al. [5] counseled that fabulous strategies for information retrieval are these constructing on the simple data retrieval methods, however extending these to permit for well-motivated compound terms and similar descriptive units. The required NLP technological know-how is now being established, and work on applying it to textual content retrieval is beginning. The traditional data retrieval services continue to make heavy use of strongly managed indexing languages, growing use is being made of indexing where phrases are drawn from the herbal language of documents.

According to Rini John et al. [6] internet entity extraction plays an essential role in Natural Language processing specifically in the data retrieval systems. A range of emerging web entity extraction techniques has been surveyed. Most of the surveyed literatures were DOM tree and database wrappers based methods. Even although they achieve right performance, they fail if the internet pages do not comply with W3C standards. It also does now not characteristic with a internet page that is complex. The issue of context switching comes with techniques primarily based on the database wrapper. In order to produce better results, however, little work has been completed towards incorporating internet entity extraction and NLP techniques.

2.4 Summarization

According to Matthew Mayo,[7] KDnuggets said that The bona fide semantic appreciation of human language text, exhibited its high quality summarization, may additionally properly be the holy grail of herbal language processing.

That announcement is not as hyperbolic as it sounds: as proper human language understanding definitely is the holy grail of NLP, and actual advantageous summarization of stated human language would necessarily entail genuine understanding.

Dr.MichaelJ.Garbade [8] said that Text summarization refers to the technique of shortening lengthy pieces of text. The intention is to create a coherent and fluent precis having only the predominant factors outlined in the document.
Automatic textual content summarization is a common hassle in machine studying and herbal language processing (NLP).

According to PavanVadapalli said that the field of textual content summarization is experiencing fast growth, and specialized tools are being developed to handle extra centered summarization tasks.

Automatic text summarization is a tool that permits a quantum bounce in human productiveness by means of simplifying the sheer quantity of facts that human beings engage with daily. This not solely allows human beings to reduce down on the studying vital however additionally frees up time to examine and apprehend in any other case ignored written works.

2.5 Dialogue Systems

According to NisarShah[9] Spoken Dialogue Systems (SDS) have been the most distinguished component in today's virtual non-public assistants (VPAs). Among these VPAs, Microsoft's Cortana, Apple's Siri, Amazon Alexa, Google Assistant, and Facebook's M, have included SDS modules in a number of devices, which permit users to talk naturally.

According to Hongshen Chen Researchers investigated on making use of neural networks to the special factors of a common task-oriented communicate system, inclusive of natural language understanding, natural language generation, talk state tracking. Recent years, end-to-end frameworks end up famous in no longer only
the non-task-oriented chit-chat communicate systems, but also the task-oriented ones.

PavloKushnery okay said that the improvement of dialogue structures is a very vital scientific and sensible venture at the moment. Improving existing and developing new techniques for building, education and checking out dialogues will assist develop in-depth studying on the one hand and improve offerings on the other.

2.6 Information Extraction (IE)

According to JiansongZhang[10] the data extraction results are summarized When calculating the precision and recall for "subject restriction" and "quantity restriction" instances, the correctness of extracting one limit instance is calculated as a ratio of the number of efficiently extracted concepts and relations to the whole number of standards and relations in that restrict (since every restriction instance may consist of a couple of ideas and relations).

ANIRUDDHA BHANDARI, J said that Text records

consists of a lot of records however no longer all of it will be necessary to you. We may be looking for names of entities, others would choose to extract precise relationships between those entities. Our intentions differ according to our requirements.

According to GerganaPetkova[11] Information extraction is the system of extracting specific (pre-specified) data from textual sources. One of the most trivial examples is when your e mail extracts solely the facts from the message for you to add in your Calendar.

3. Research Findings

Since we use language to interact with our devices, NLP became an integral part of our lives. While considering the different applications of NLP, for each of them there are lots of frame works and methods available now. We cannot completely rely on a particular method because, each method has its own advantages and disadvantages. So we need to focus on the integration of a variety of methods together to achieve high accuracy and quality for the applications like machine translation, information retrieval etc..And in the field of linguistic knowledge we need to do more researches. So it can conclude that usage of NLP help us to interact with the technical world more easier, still now research are going based on the technology, hence NLP will useful in the coming future.

4. CONCLUSIONS

While NLP is a noticeably latest location of appear to be up and application, as in contrast to exceptional data technological information approaches, there have been sufficient successes to date that advocate that NLP-based facts get entry to utilized sciences will proceed to be a crucial vicinity of appear up and enchancment in files constructions now and a prolonged way into the future.

References

[1] Neeha Ashraf, Manzoor Ahmad, Machine Translation Techniques and their Comparative Study, International JournalofComputerApplications (0975 – 8887),Volume 125 – No.7, September 2015

[2] Li Peng, A Survey of Machine Translation Methods ,TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7125~7130 e-ISSN: 2087-278X

[3] A.Clementeena , Dr. P. Sripriya ,A literature survey on question answering system in natural language processing, International Journal of Engineering & Technology, 7 (2.33) (2018) 452-455

[4] O.Ferret, B.Grau, M.Hurault-Plantet, G.Illouz, C.Jacquemin, L.Monceaux,I.Robba, A.Vilnat,How NLP can improve Question Answering ,Article in Knowledge Organization · January 2002

[5] David D. Lewis, Karen SparckJones,Natural language processing for information retrieval, Communications of the ACM, 39 (1), 1996, 92-101.

[6] Rini John, SharvariGovilkar, INFORMATION RETRIEVAL TECHNIQUE FOR WEB USING

NLP,International Journal on Natural Language Computing (IJNLC) Vol. 6, No.5, October 2017

[7] PRATEEK JOSHI, NOVEMBER 1,2018, An Introduction to Text Summarization using the TextRank Algorithm

[8] Dr. Michael J. Garbade Sep 19, 2018 A Quick Introduction to Text Summarization in Machine Learning

[9] Nisar Shah Jul 23, 2018 Introduction to Dialogue Systems.

[10] Zhang, J. and El-Gohary, N. (2015). "Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking." J. Comput. Civ. Eng., 10.1061/(ASCE)CP.1943-5487.0000346, 04015014.

Author Profile

Shanet Mathew received Degree in Bachelor of Computer Applications from Mahatma Gandhi University Kottayam, Kerala in 2017.

Shilpa CS received Degree in Bachelor of Computer Applications from Kerala UniversityTrivandrum,Kerala in 2020.

Dr. Bijimol TK has been working as an Assistant Professor in the department of Computer Science sine June 2002. She has completed her PhD from Bharathiar University Coimbatore, Tamil Nadu. She has published a book and has 10 more journal and proceedings publications.