Natural Language Processing: Opportunities in Information Technology

Deepmala A. Sharma*1

Asst. Prof., Department Of Computer Science
Tilak Maharashtra Vidyapeeth, Gultekadi,
Pune 411 037, Maharashtra, India.

Abstract

The basic purpose of this research is to highlight technological issues in human and machine communication in terms of voice and speech. This discussion is further explained with the major issues in implementing human and machine communication. Also this study reflects the rise in job ration in the field. Natural language processing (NLP) is a branch of computer science that focuses on development of machines in terms of communication with humans in their native language. NLP is the area of study dedicated to text and speech automation. It is an old field of study, originally subjected using by rule-based methods designed by linguists, then statistical methods with machine learning, and, more recently, deep learning methods that shows great promise in the field. However due to language ambiguity the system is found inefficient in performing certain tasks. But the future of NLP is bright with the emergence of big data and data science has helped in several decisions driven task making. NLP for Big Data is the Next Big Thing. NLP can solve big problems of the business world by using Big Data. Be it any business like retail, healthcare, business, financial institutions [18], education, industry the use of NLP can be noted. There is a great deal of job opportunities for IT with knowledge of data science, machine learning and big data. In the coming years, machines are expected to do lot more than they are doing presently with applications such as NLP at the core. As per the latest report published by Market Research Future (MRFR), the global market for NLP is projected to surge at 24% CAGR during the assessment period (2017-2023) [13].

Keywords: Human, Machine, Artificial Intelligence, Natural Language Processing, machine learning, data science, ambiguity, automation, deep learning, Big Data.

Introduction

In the current trends of technology where language plays a major role in communication Natural language processing (NLP) can be defined as the ability of machine to analyze, understand and generate machine [2] that can process human speech or textual information. NLP deals with understanding native languages of peoples. The process includes machine learning and optimization, statistical modeling, linguistic and cognitive approach. The aim of NLP is to make computers interact with humans exactly the way humans communicate with each other. The means of communication in NLP could be through typed text or spoken words. However human language is available in diversity which makes it difficult to process by machines. NLP is extensively used for text mining, machine translation, and automated question answering [1]. While computers have been found inefficient in performing certain process of NL efficiently, the main reason is that machines work efficiently on instructions given in structured form using programming languages. Whereas natural language is a language full with unstructured data as it deals with native languages. The data availability in structured format is limited and majority of data is available in textual form which is highly unstructured in nature. In order to produce significant and actionable insights from this data it is important to get acquainted with the techniques of text analysis and natural language processing. Therefore analyzing, recognizing and converting unstructured data is a big concern in NLP. However continuous research and development can be observed in NLP from past years. In the last few years NLP with the use of machine learning and deep analysis techniques for information processing there has been tremendous improvement found in NLP devices. Devices like Google assistant, Apple Siri, Amazon Alexa are the most popular and powerful voice assistant incorporated in smart devices now a days.
Natural language processing (NLP) helps in converting unstructured data into structured format. But the truth is human language is highly ambiguous. It is also ever changing and evolving. Humans are great at producing language and understanding language, and are capable of expressing, perceiving, and interpreting in a detailed manner. But, for a machine it is difficult and challenging task to process unstructured data and react in natural form as it is has artificial mind.

Text mining and text analytics is the process of deriving meaningful information from natural language. It usually involves structuring the input text, deriving patterns of structures in available data and finally interpreting the output.

Applications of NLP
1. Sentimental analysis
2. Chat board
3. Speech recognition (voice assistant: Google assistant, Seri and Quartana)
4. Machine translation (to translate data from one language to another)
5. Another application includes – spell check, keyword search, information extraction, advertisement matching, automatic summarization, data extraction, Plagiarism detection, automatic translation and more.

However till date the systems available lack in accuracy and clarity in speech recognition and continuous research and development is being carried out for the same. The reasons is there is a need of improvement in terms of methodology and technology that helps in better speech recognition and speech translation of spoken language into text by computers. The ultimate goal of NLP is to fill the gap how the humans communicate (natural language) and what the computer understands (machine language) [18].

The main problem with NLP processing is that the machines follows semantic approach Human language is rarely spoken in a structured way. So, in order understand human language is to understand not only the words, but the concepts and how they’re linked together to create meaning [1] is also necessary. Despite language being one of the easiest things for humans to learn, the ambiguity of language is what makes natural language processing a difficult problem for computers to master [1].

To overcome the above problems NLP technologies today are powered by Deep Learning — a subfield of machine learning. Deep Learning only started to gain momentum again at the beginning of this decade. Deep Learning provides a very flexible, universal, and learnable framework for representing the world, for both visual and linguistic information.

Initially, it resulted in breakthroughs in fields such as speech recognition and computer vision. Recently, deep learning approaches have obtained very high performance across many different NLP tasks.
Results

Above we have tried to understand the AI technology using Neural Engine in devices which is a big innovation in machine language technology. However every technology has some scope of upgradation such as:-

1. Apple uses Siri as a virtual assistant that is part of Apple. The assistant uses voice queries and a natural-language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of Internet services [5] but country like India where there are 22 major languages written in 13 different scripts, with over 720 dialects it would be very difficult to use this technology in day to day business process due to availability of wide range of vocal accent available as it requires.

2. Similarly Google uses AI in Google Assistant to assist voice queries and a user interface to answer our questions also recently a new tool calledGoogle Indic keyboard allows you to speak in your native Indian language and translates spoken words in written form which can be used in socialmedia as a communication tool. E.g. What’s app. But it limits the communication if words are not expressed in a defined Scope. We describe the past development of NLP, and summarize common NLP sub-problems in this broadfield [15].

With reference to result we can state that there is a lot of scope for improvement in NLP and this can be achieved with deep learning technique which deals with study of neural network method which is based on machine learning.

Research Methodology

This above research is based on secondary data collected from previous research papers and web sources. The proposed study focuses on NLP techniques and variety of methods that enables a machine to understand what’s being said or written in human language. Currently Natural language processing is based on deep learning methodology. It tries to find out relation between samples of data and incorporate them together into the desired outcome. The algorithm uses descriptive and predictive analysis for the outcome to result set.

Conclusion:

The field of natural language processing is shifting from statistical methods to neural network methods. There are lot more challenging problems to solve in natural language. However, deep learning methods are achieving state-of-the-art results on specific language problems [6]. In terms of Information Technology (IT) Natural language processing has gained a great deal of demand rising in permanent job opportunities.

References

[3] https://heartbeat.fritz.ai/the-7-nlp-techniques-that-will-change-how-you-communicate-in-the-future-part-i-f0114b2f0497
[9] https://towardsdatascience.com/an-easy-introduction-to-natural-language-processing-b1e2801291c1
[10] https://web2.qatar.cmu.edu/~egakhar1/NLP.html
[12] https://www.itjobswatch.co.uk/jobs/uk/natural%20language%20processing
[16] https://www.slideshare.net/jaganadhg/natural-language-processing-14660849?next_slideshow=1