

Current Trends in Natural Language Processing

Abhishek Mehta¹, Dr. Kamini Solanki²

¹Assistant Professor (Parul Institute of Computer Application, Parul University, India)

²Associate Professor (Parul Institute of Computer Application, Parul University, India)

Abstract—Natural language processing is a branch of computer science and artificial intelligence which is concerned with interaction between computers and human languages. Natural language processing is the study of mathematical and computational modelling of various aspects of language and the development of a wide range of systems. These includes the spoken language systems that integrate speech and natural language. Natural language processing has a role in computer science because many aspects of the field deal with linguistic features of computation. Natural language processing is an area of research and application that explores how computers can be used to understand and manipulates natural language text or speech to do useful things. The applications of Natural language processing include fields of study, such as machine translation, natural language text processing and summarization, user interfaces, multilingual and cross language information retrieval (CLIR), speech recognition, artificial intelligence (AI) and expert systems.

Index Terms—Natural language Processing (NLP), Cross Language Information Retrieval (CLIR), Artificial intelligence (AI)

INTRODUCTION

Natural Language processing is a branch of computer science, artificial intelligence and linguistics concerned with the interactions between computers and human (natural) language. Natural languages are languages spoken by humans. Natural language is any language that humans learn from their environment and use to communicate with each other. Whatever the form of the communication, natural languages are used to express our knowledge and emotions and to convey our responses to other people and to our surroundings. Natural languages are usually learned in early childhood from those around us. Currently we are not yet at the point where these languages in all of their unprocessed forms can be understood by computers. Natural language processing is the collection of techniques employed to try and accomplish that goal. The field of natural language processing (NLP) is deep and diverse. Natural language processing (NLP) is a collection of techniques used to extract grammatical structure and meaning from input in order to perform a useful task as a result, natural language generation builds output based on the rules of the target language and the task at hand. NLP is useful in the tutoring systems, duplicate detection, computer supported instruction and database interface fields as it provides a pathway for increased interactivity and productivity.

literature review

The research work in the natural language processing has been increasingly addressed in the recent years. The natural language processing is the computerized approach to analyzing text and being a very active area of research and development. The literature distinguishes the main

application of natural language processing and the methods to describe it.

Natural language processing for Speech Synthesis:

This is based on the text to speech conversion i.e (TTS) in which the text data is the first input into the system. It uses high level modules for speech synthesis. It uses the sentence segmentation which deals with punctuation marks with a simple decision tree.

Natural language processing for Speech Recognition:

Automatic speech recognition system make use of natural language processing techniques based on grammars. It uses the context free grammars for representing syntax of that language presents a means of dealing with spontaneous through the spotlighting addition of automatic summarization including indexing, which extracts the gist of the speech transcriptions in order to deal with Information retrieval and dialogue system issues.

levels of NLP

The most explanatory method for presenting what actually happens within a Natural Language Processing system is by means of the 'levels of language' approach. This is also referred to as the synchronic model of language and is distinguished from the earlier sequential model, which hypothesizes that the levels of human language processing follow one another in a strictly sequential manner. Psycholinguistic research suggests that language processing is much more dynamic, as the levels can interact in a variety of orders. Introspection reveals that we frequently use information we gain from what is typically thought of as a higher level of processing to assist in a lower level of analysis. For example, the pragmatic knowledge that the document you are reading is about biology will be used when a particular word that has several possible senses is encountered, and the word will be interpreted as having the biology sense. Of necessity, the following description of levels will be presented sequentially. The key point here is that meaning is conveyed by each and every level of language and that since humans have been shown to use all levels of language to gain understanding, the more capable an NLP system is, the more levels of language it will utilize.

Phonology:

This level deals with the interpretation of speech sounds within and across words. There are, in fact, three types of rules used in phonological analysis [11]:

- 1) Phonetic rules: It is used for sound within words.
- 2) Phonemic rules : It is used for variations of pronunciation when words are spoken together.
- 3) Prosodic rules : It is used to check for fluctuation in stress and intonation across a sentence.

In an NLP system that accepts spoken input, the sound waves are analyzed and encoded into a digitized signal for

interpretation by various rules or by comparison to the particular language model being utilized.

Morphology:

Morphology is the first stage of analysis once input has been received. It looks at the ways in which words break down into their components and how that affects their grammatical status. Morphology is mainly useful for identifying the parts of speech in a sentence and words that interact together. The following quote from Forsberg gives a little background on the field of morphology. Morphology is a systematic description of words in a natural language. It describes a set of relations between words' surface forms and lexical forms. A word's surface form is its graphical or spoken form, and the lexical form is an analysis of the word into its lemma (also known as its dictionary form) and its grammatical description. This task is more precisely called inflectional morphology. Being able to identify the part of speech is essential to identifying the grammatical context a word belongs to. In English, regular verbs have a ground form with a limited set of modifications, however, irregular verbs do not follow these modification rules, and greatly increase the complexity of a language. The information gathered at the morphological stage prepares the data for the syntactical stage which looks more directly at the target language's grammatical structure.

1) Syntax: Syntax involves applying the rules of the target language's grammar, its task is to determine the role of each word in a sentence and organize this data into a structure that is more easily manipulated for further analysis. Semantics are the examination of the meaning of words and sentences.

a) Grammar: In English, a statement consists of a noun phrase, a verb phrase, and in some cases, a prepositional phrase. A noun phrase represents a subject that can be summarized or identified by a noun. This phrase may have articles and adjectives and/or an embedded verb phrase as well as the noun itself. A verb phrase represents an action and may include an imbedded noun phrase along with the verb. A prepositional phrase describes a noun or verb in the sentence. The majority of natural languages are made up of a number of parts of speech mainly: verbs, nouns, adjectives, adverbs, conjunctions, pronouns and articles.

b) Parsing: Parsing is the process of converting a sentence into a tree that represents the sentence's syntactic structure. The statement: "The green book is sitting on the desk" consists of the noun phrase: "The green book" and the verb phrase: "is sitting on the desk." The sentence tree would start at the sentence level and break it down into the noun and verb phrase. It would then label the articles, the adjectives and the nouns. Parsing determines whether a sentence is valid in relation to the language's grammar rules.

C. Semantics: It builds up a representation of the objects and actions that a sentence is describing and includes the details provided by adjectives, adverbs and propositions. This process gathers information vital to the pragmatic analysis in order to determine which meaning was intended by the user.

D. Pragmatics: Pragmatics is "the analysis of the real meaning of an utterance in a human language, by disambiguating and contextualizing the utterance". This is accomplished by identifying ambiguities encountered by the

system and resolving them using one or more types of disambiguation techniques .1) Ambiguity: Ambiguity is explained as "the problem that an utterance in a human language can have more than one possible meaning.

applications of nlp application of natural language processing:

Autocorrect and Autocomplete

Chatbots & Virtual Assistants

Speech Recognition and Voice Assistants

Targeted Advertising

Machine Translation and Language Translator

Auto-correct and Auto-complete

Whenever you search anything on Google, after typing 3-4 letters, it shows you the possible search terms. Or, if you search for something with types, it corrects them and still shows relevant results for you. Isn't it amazing?

It is something that every person uses this in daily life but never pays much attention on it. It's a wonderful application of natural language processing and a great example of how it's affecting millions of people around the world, including you and me. auto-complete and auto-correct both help us in finding accurate search results much efficiently. Now a day's various other companies have also started using this feature on their websites like Google, Facebook, Instagram and Quora.

Chatbots & Virtual Assistants

Chatbots & virtual assistants are used for automatic queries solving, designed to understand natural language and deliver an appropriate response through natural language generation. Customer care is the most important thing for any company. It can help the company's improve their products, and also keep the users satisfied. But get across with every customer manually, and helping them for their problems can be an exhausting task. that's why Chatbots come into the picture. Chatbots help the company's in achieving the goal of smooth customer experience.

Today, many companies use chatbots for their applications and websites, which solves basic questions of a user's. It not only makes the process easier for the company's but also saves customers from the frustration of waiting to interact with customer call assistance.

It can cut down the cost of hiring call center agent for the company. Initially chatbots were only used as a tool that deal with customer's quires, but nowadays they have evolved into a personal companion. From recommending a product to taking feedback from the users, chatbots can do everything.

Speech Recognition and Voice Assistants

Speech recognition technology uses natural language processing to transform spoken language into a machine-understandable format. you have already met them Google Assistant, Apple Siri, Amazon Alexa. All of these are voice assistants.

A voice assistant is a module that uses for speech recognition. natural language processing to understand the voice commands of a user and perform actions accordingly. It is like to a chatbot, but I have included voice assistants independently because they deserve a better place on this list.

They are not like a chatbot and can do many more things than a chatbot.

Nowadays, most of us cannot imagine our lives without voice assistants. Throughout the years, they have revolutionized into a very reliable and powerful friend. From setting our alarm's to finding a hotel for us, a voice assistant can do anything. They have opened a new path of opportunities for both users and companies.

Targeted Advertising

Whenever you searching for anything on Amazon, and a few minutes later, Google started showing you ads related to similar product on various webpages.

Targeted advertising is a type of online advertising where ads are shown to the user based on their online searches. Most of the online companies today use this system because first, it saves companies a lot of money, and second, relevant ads are shown only to the potential customers.

Targeted advertising works mainly on Keyword Matching. The Advertisement are associated with a keyword or wording, and it is shown to only those users who search for the keyword similar to the keyword with which the advertisement was associated. Obviously, not only that, there are other factors like the websites they visited recently, and the webpages they showed interest in, are all taken into account to provide the users with the relevant ads of products that they might be searching.

Machine Translation and Language Translator

Google Translate to find out what a particular word or phrase is in a different language and the ease with which it translates a piece of text in one language to another is pretty amazing, the technique behind it is Machine Translation. Machine translation is one of the first applications of natural language processing. In older days, machine translation systems were dictionary based and rule based system, and they got limited success. However, due to change in generation in the field of neural networks, availability of humongous data, and powerful machines, machine translation has become fairly accurate in converting the text from one language to another. Nowadays, tools like Google Translate can easily convert text from one language to another language. These tools are helping various people and businesses in breaking the language barrier and becoming successful.

ADVANTAGES

Users can ask questions about any subject and get direct answers within seconds. NLP system provides response to the questions in natural language. NLP system provides exact answers to the questions, there is no unnecessary or unwanted information's. The accuracy of the answers increases with the amount of relevant information given in the question. NLP process helps computers convey with mans in their language and scales to other language and related tasks. Allows you to presents more language-based data compares to a human being without fatigue and in an unbiased and consistent way. Structuring a highly irregular data source Produce a readable summary of a part of the text.

Given a sentence or larger chunk of text, determine comparable words refer to the same objects. This includes a number of similar tasks. One task is identified the discourse structure of connected text

DISADVANTAGES

Complex Query Language- the system may not be able to give the correct answer it the question that is poorly worded or ambiguous. The system is construct for a single and specific task only; it is unable to adapt to new domains and problems because of limited functions. NLP system doesn't have a user interface which lacks features that allow users to further communicate with the system.

CONCLUSION

While NLP is a relatively recent area of research and application, as compared to other information technology approaches, there have been sufficient successes to date that suggest that NLP-based information access technologies will continue to be a major area of research and development in information systems now and far into the future. The state-of-the-art Natural Language Processing techniques applied to speech technologies, specifically to Text-To-Speech synthesis and Automatic Speech Recognition. In 3TTS. The importance of NLP in processing the input text to be synthesized is reflected. The naturalness of the speech utterances produced by the signal-processing modules are tightly bound to the performance of the previous text-processing modules. In ASR the use of NLP particularly is complementary [7]. It simplifies the recognition task by assuming that the input speech utterances must be produced according to a predefined set of grammatical rules. Its capabilities can though be enhanced through the usage of NLP aiming at more natural interfaces with a certain degree of knowledge. Reviews the major approaches proposed in language model adaptation in order to profit from this specific knowledge

REFERENCES

- [1] Acharya, T.K.T. (1973). "Scope of Green Revolution in Maharashtra", State Bank of India, Monthly Review, Vol. 12, No.10, pp. 369-386.
- [2] Acharya, Shabd S. (2007), "National Food Polices Impacting on Food Security: The Experience of a Large Populated Country-India", in BasudeoGuhaKhasnosis, Shabd S. Acharya and Benjamin Davis (Eds) Food Insecurity, Vulnerability and Human Rights Failure, Studies in Development Economics and Policy, World Institute for Development Economics.
- [3] Research. AERC (1994) , " Inter District Variations in Agricultural Development in Gujarat-1949-50 to 1991-92".M. T. Pathak , H. F. Patel and R. A. Dutta, Agro Economic Research Centre, Vallabh Vidyanagar.
- [4] AERC (1994), "Inter district Variation in Agriculture Development in Rajasthan1956-557 to 1989-90", D. Bagchi and Verma, H.M., Agro Economic Research Centre, Vallabh Vidyanagar.
- [5] Agrawal A.N., "Indian Economy Problem of Development and Planning", 16th edition, Wiley Eastern Ltd., N. Delhi, 1990.
- [6] Bagchi D., "Review of state Agriculture Policy in Rajasthan", AERC, SPU, VVNagar, 2006.
- [7] BagchiDilip (2004), "Review of State Agricultural Policy in Rajasthan", Agroeconomic Research Centre, Vallabh Vidyanagar, Gujarat State.
- [8] Ball, V. Eldon; Bureau, Jean-Christophe; Nehring, Richard and Somwaru, Agapi (1997), "Agricultural productivity Revisited" American Journal of Agricultural Economics, Vol.79, No. 4, November, pp. 1045-1063.
- [9] Bhalla G.S. and Alagh Y.K., "Performance of Indian Agriculture, A District wise Study", Sterling publishers pvt.ltd. N. Delhi, 1980. 379 10. Bhalla G.S. and Singh Gur mail, "Economic Liberalizati