

Natural Language Processing

Riddhi Chauhan, Mahesha Sisodiya , Nirali Suthar

Student (IT department , Grow more Faculty of engineering , himantnager , Gujarat , India)

Student (IT department , Grow more Faculty of engineering , himantnager , Gujarat , India)

Student (IT department , Grow more Faculty of engineering , himantnager , Gujarat , India)

Date of Submission: 08-12-2022

Date of Acceptance: 16-12-2022

ABSTRACT—The method incorporates a language like English has invariably been one of the central analysis problems with computer science, each thanks to the key role language plays thanks to human intelligence and potential wealth Applications will use language process techniques to build it attainable to use language to precise assume programming, this flip will increase accessibility Programming for non-expert users. informatics holds nice promise for creating laptop interfaces easier to use for individuals. Informatics is employed to investigate text, then permitting machines to know however humans speak. During this paper, I tend to provide a summary of informatics from scratch. tend to additionally , in short, discuss a number of its major applications

Keywords – NLP(Natural Language Processing), ML(Machine Learning), SR(Speech Recognition), NLU(Natural Language Understanding), NLG(Natural Language Generation)

I. INTRODUCTION

The Natural language process is a locality and associate application that explores , however, computers are used to perceive and manipulate language text or speech to try to do helpful things. informatics is multi-disciplinary, it's closely similar to humanities. It additionally has links to computers and Information Science, Psychology, electrical and Electronic Engineering is, of course, additionally associated with unnatural intelligence

Applications - MT, speech recognition, text method , and outline than on.

DESCRIPTION

Natural Language process could be a field of technology, computer science , and procedure humanities that deals with the study of interactions between laptops and Human's Natural Languages. In alternative words, a way to perceive teaching/programming computers and Generate human language.

Natural Language process could be a term of exploration and application that explores however laptop is accustomed be perceive and handle language text or speech to try to helpful things language process will be outlined because the automatic process of human language. over one field what's typically chosen is, 'Language Technology' or 'Language Engineering. it's closely analogous to humanities.

To summarize, informatics could be a discipline that's relative to natural human languages and interaction between Computing devices.

It is some way in which computers analyze, perceive and obtain means from good , and helpful human language means this human-computer interaction permits the \$64000 world Applications like a computer program, translation Systems, machine-driven Question respondents, Text classification, writing systems , and descriptive linguistics checking and a lot of informatics is wide integrated into giant numbers References like assessment systems, e-learning, research, MT, trilingual and cross-language info Retrieval, Speech Recognition.

As computers play a significant role in the preparation, storage,

analysis and transfer of knowledge, endowing them with the ability to grasp and get info expressed in language becomes necessary. Hence, the goal of IP is to style and build laptop systems that have the power to investigate natural languages (like English, German, etc.) and to get output in natural language. Human communication happens as speech and as written communication. NLP's language process downside will be divided into two chores:

1) process the transcription, this may be done by absorption of lexical, syntactical, and linguistic data of the language.

2) process speech, this may be done absorption of all the information required and extra data regarding phonology.

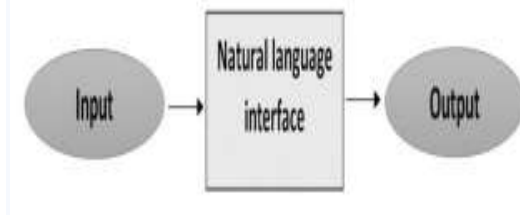
NLP system includes:

- 1) User input
- 2) It goes to the tongue interface
- 3) Output obtained during a lang

The input associated degree output of a human language technology system are often of 2 types:

A. parts of human language technology

There are 2 parts of human language technology.



1. Language Understanding (NLU): It involves mapping the input into helpful representation and analyzing completely different aspects of the language
2. Language Generation (NLG): It involves producing purposeful sentences in language from the representation.

B. human language technology language

1. Phonology: Study of however sounds area unit organized and used in natural languages
 2. Morphology: Study of words, however, they're shaped ,and their relationship to alternative words within the same language.
 3. Syntax: Arrangement of words and phrases to form well-formed sentences during a language.
 4. Semantics: Study of meanings of words and phrases during a language. Has 2 main areas: lexical linguistics and logical linguistics
 5. Pragmatics: Deals with mistreatment and understanding sentences in numerous things and the way the interpretation of the sentence is affected.
- C. Steps in human language technology There area unit typically 5 steps:

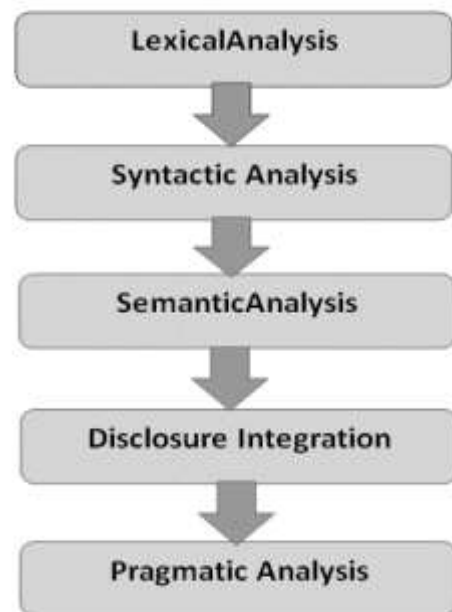
C. Steps in human language technology

Their area unit typically 5 steps:

- A. Lexical analysis: It involves distinguishing ,and analyzing the structure of words. it's essentially dividing the total chunk of text into paragraphs, sentences and words.
- B. Syntactical analysis (parsing): This involves the analysis of words within the sentence for descriptive linguistics and transcription words to show the connection among the words.
- C. Linguistics analysis: It attracts the precise means or the dictionary which means of the

text. The text is checked for meaningfulness by mapping syntactical structures and objects within the task domain.

- D. Discourse integration: the which means of any sentence depends upon the which means of the sentence preceding and at once succeeding it
- E. Pragmatic analysis: it involves etymologizing those aspects of language that need planet information. During this, what was aforementioned is re-interpreted on what it meant.



II. BACKGROUND

NLP is unbelievably precious. it's over fifty years of history as a scientific discipline. The history of human language technology usually started in the 1950s, though work is often found from earlier periods.

Early add human language technology was targeted at a generation of complete computer programs that might compile and run. The dictionary look-up system developed in 1948 was the primary recognizable human language technology application. In 1949, Warren Weaver was involved in code-breaking throughout the Second warfare.

During the Fifties, most human language technology researchers were concentrating on computational linguistics (Russian to English) as AN application. Turing printed a writing 'Computing Machinery and Intelligence' that projected the Turing Test. Alan Turing's check could be a check of a machine's ability to exhibit intelligent behavior indistinguishable from that of a

person. In 1957, a linguistic scientist, a young Yankee Linguist, introduced the thought of generative synchronic linguistics rule-based description of syntactic structures. Most of human language technology since then, has been marked by his influence.

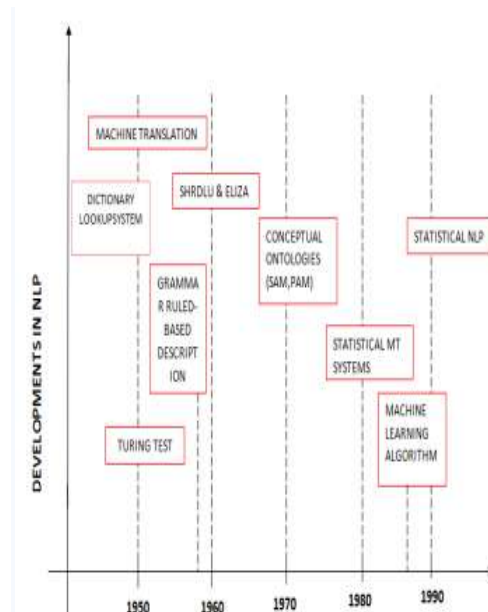
In the mid-60s computational linguistics worked solely word-by-word and funding got drastically reduced. In the 1960s SHRDLU and ELIZA were some of the human language technology systems developed.

During the Seventies, several programmers began to write down “conceptual ontologies (SAM, PAM)” and structured real-world information into machine-graspable knowledge.

Up to the Nineteen Eighties, human language technology systems supported complicated sets of handwritten rules. However, in the late Nineteen Eighties, there was a revolution in natural language processing (NLP) human language technology (information science/informatics/information processing/IP) with the introduction of machine learning algorithms like call trees for language processing. The first applied mathematics computational linguistics systems were developed in the Nineteen Eighties.

And there have been conjointly some key developments like increased Transition Networks, cost grammar, and linguistics representations.

In the Nineties applied mathematics and human language technology became the foremost common paradigm. Recent analysis has targeted unattended and semi-supervised learning algorithms. Such algorithms are ready to learn from the knowledge that has not been navel-annotated with the desired answers. By 2025, trade consultants expect human language technology to be ready to method all human languages at AN accuracy of 99%.



NLP ARCHITECTURE

We need the tongue as an associated input for human language technology and to urge the tongue as an output we'd like the human language technology system. The NLP system has 2 components, understanding half (input facet processing) and generating half (output facet processing). The general human language technology system is shown in fig.

The human language technology system includes:

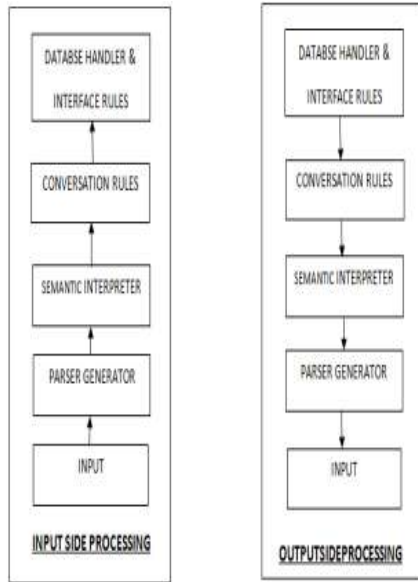
1. A computer program
2. linguistics Interpreter
3. Conversion Rules
4. information Base Handlers

Each of the elements square measure explained as follows:

1. Parser: Once the input is given to the computer program it generalizes a grammar structure within the style of take apart tree. The parsing aim is to structure the unstructured text. The elemental plane of parsing is to cluster words together to create phrases that behave as one unit. These phrases are often combined along to create larger phrases and eventually sentences. A computer program is employed to determine whether or not a given string belongs to the language and maps a string of words to its take- apart tree.
2. Semantic Interpreter: linguistics interpreter captures the semantic details of the analyze tree and generates a deeper
3. structure of the analysis tree. Conversion Rules: the conversion rules settle for the deep

structure of a sentence from the linguistics interpreter and build it compatible to be kept within the information.

4. information Handler: the information handler works on the modified deep structure and generates a processed type for storage.



$S \rightarrow NP VP$
$VP \rightarrow V_i$
$VP \rightarrow V_i NP$
$VP \rightarrow VP PP$
$NP \rightarrow DT NN$
$NP \rightarrow NP PP$
$PP \rightarrow IN NP$

S= Sentence, VP= predicate, NP= phrase, PP= prepositional phrase, VI= verb, VT= transitive verb, DT= determiner.

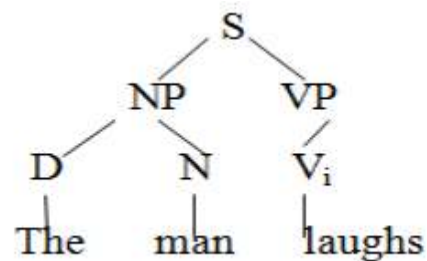
A. Leftmost Derivation

A leftmost derivation may be a sequence of strings, S_1, \dots, S_n , where $S_1=S$, the beginning image. $S_n \in \Sigma^*$, i.e.; metal is formed from only terminal symbols.

E.g.: [S], [NP VP], [D N VP], [the N VP], [the man VP], [the man V_i], [the person laughs]

Representation of derivation as a tree:

An Example: the person laughs.



CONTEXT-FREE GRAMMAR

Context-free descriptive linguistics could be an ordinarily used mathematical system for modeling syntax in Natural Language .It was 1st outlined for Natural Language in 1957 by A. Noam Chomsky .Context-Free descriptive linguistics belongs to the sphere of FLT wherever communication is viewed as a collection of sentences; a sentence as a string of words from the vocabulary of the language & grammar is finite.

So, Context-free descriptive linguistics consists of four components:

- Σ , the terminal vocabulary: the words of the language being outlined
 - N, the non-terminal vocabulary: a collection of symbols disjoint from T.
 - R, a collection of rules of the shape $X Y_1, Y_2, \dots, Y_n$
 $N_0, X \in N, Y_i \in (N \cup \Sigma)$
 - S, a start symbol, a member from N.
- Example: A Context-free grammar for English
 $N = \{S, NP, VP, PP, DT, VI, VT, NN, IN\}$

$S=S$ (Start symbol)
 $\Sigma = \{\text{words in the language}\}$
 R= Set of rules

Derivation	Rules used
S	$S \rightarrow NP VP$
NP VP	$NP \rightarrow DT N$
DT N VP	$DT \rightarrow \text{the}$
The N VP	$N \rightarrow \text{man}$
The man VP	$VP \rightarrow V_i$
The man V_i	$V_i \rightarrow \text{laughs}$

In the derivation, we tend to begin with the beginning image 'S', and in each step we have

a tendency to outline a rule for the derivation. i.e.,
 S
 goes to NP VP (S NP VP).

And then, within the next step, we tend to replace 'S' within the derivation with NP VP. the fundamental plan here is that at each step, we tend to choose the left most non-terminal image in the derivation & replace it with the non-terminals defined within the rules, and it's continuing to urge a

sequence of words. therefore an entire derivation forever ends with a string wherever each word in a very strong could be a word within the language outlined. Then finally, these derivations are a unit pictured as a take apart tree.

So, Context-Free descriptive linguistics primarily defines a group of possible derivations

ENGLISH SYNTAX

A. Part of Spoken

Tags	Types	Examples
NN	Singular noun	Man, dog, park.
NNS	Plural noun	Flowers, houses, cars.
NND	Proper noun	Sita, IBM.

01. Nouns:
02. Determiners: Determiners come before nouns.

Tag	Examples
DT	The, a, some, every.

03. Adjective: Adjectives usually come between determiners and nouns.

Tag	Examples
JJ	red, green, large, small etc.

B. Noun Phrase Grammar

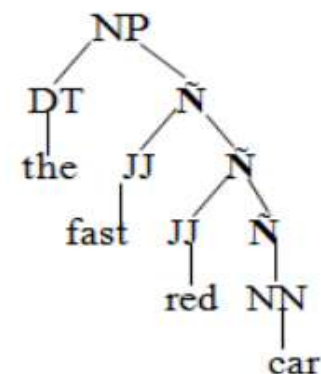
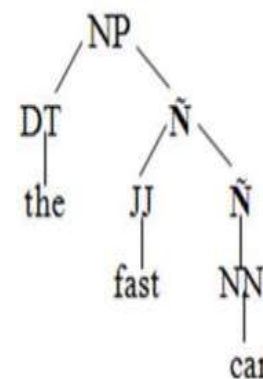
A noun phrase is formed by a determiner (DT)

followed by noun

Rules:

$NP \Rightarrow DT \bar{N}$
$\bar{N} \Rightarrow NN$
$\bar{N} \Rightarrow NN \bar{N}$
$\bar{N} \Rightarrow JJ \bar{N}$
$\bar{N} \Rightarrow \bar{N} \bar{N}$

Example:



PARSING

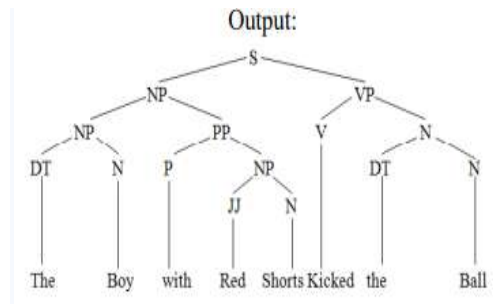
Parsing could be a technique in NLP that checks if the sentence is correct according in the synchronic linguistics and if therefore returns a analysis tree representing the structure of the sentence. So, an analyzed tree could be a tree structure with the words of sentence because of the leaves of the tree.

The elemental plane of parsing is that words cluster together to make phrases that behave as one unit. These phrases will mix along to make larger phrases and eventually sentences.

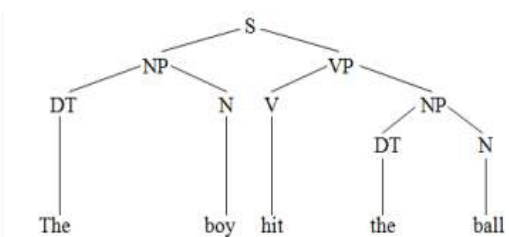
And essentially, a programme is employed to see whether or not a given string belongs to the

language and maps a string of words to its analysis tree.

S: Sentence
 NP: phrase
 VP: predicate
 DT: Determiner
 JJ: Adjective

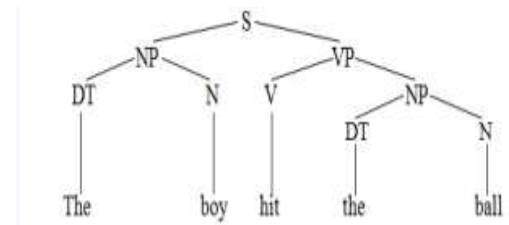


The information represented by the parse trees:
 A. Parts of the speech for each word.
 N = noun
 V = verb
 DT = determiners

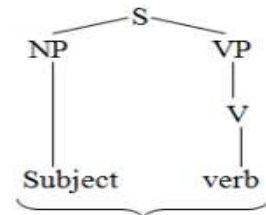
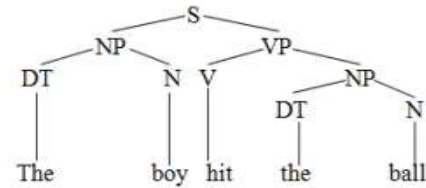


The first level of the tree encodes the parts of speech tag sequence for the input sentence.

B. Phrases At the next level, there is a hierarchical grouping of words into phrases (constituents)



Noun phrases (NP): "The boy", "the ball"
 Verb phrases (VP): "hit the ball"
 Sentences(S): "The boy hit the ball"
 C. Grammatical Relationships



Subject-verb relationship

"The boy" is the subject of "hit". These grammatical relationships allow us to identify who did what to whom.

APPLICATIONS OF NLP

Humans perform the majority of activities through language either by direct communication or by mistreatment of natural language. So, to speak, there's a necessity to grasp the language we have a tendency to use. Because the study of human languages developed, the concept of an act with no-human devices was investigated. This idea gave rise to the Natural language process. Thus, NLP helps machines to browse text by simulating the human ability to grasp the language. We use NLP applications every day and all over directly or indirectly although we have a tendency to not know it.

The applications of NLP will be divided into 2 classes as follows:

1) Text-based applications:

As the name says, text-primarily based applications are applications that traumatize the process of written communication like books, newspapers, e-mail messages, manuals, reports, research papers and lots of a lot of. These applications like extracting info from messages, articles, documents, machine translation and summarizing texts then on.

2) Dialogue-based applications:

Dialogue primarily-based applications traumatize speech communication. This involves applications like question-answering systems, tutoring systems, interactive downside determination, voice-controlled machines and

machine-driven client service over the telephone.

A. Machine Translation:

Machine translation is that the automatic translation of text or speech from one language to another. It's the matter of converting a supply text to a different language by victimization automated computing while not human intervention. Developed within the Fifties, artificial intelligence is additionally referred to as automatic translation. Today, an oversized quantity of knowledge is generated on-line and also the need to access it becomes important. Different countries have totally {different|completely different} languages and different cultures. Hence, there's a necessity of inter-language translation to share ideas, to transfer data and communicate with one another. Machine translation helps North American nations conquer language barriers that we often encounter by translating technical manuals, content or catalogs at a considerably reduced price. It offers a mere scalable variety to harmonize the world's data.

There are unit 3 sorts of machine translation:

- a) Rules based mostly systems: They use a mix of grammatical and language rules and dictionaries for words.
- b) applied mathematics systems: They analyze the big quantity of knowledge for each language combined and learn to translate.
- c) Neural systems: They incorporate an Associate in Nursing approach that makes the machines learn to translate the text through a large neural network.

B. Fighting spam:

With the recognition of the web, e-mail has become a part of our everyday life. usually emails are a unit liable to spam emails. Spam emails are unit unsought bulk email. These emails advertise merchandise and services or they cause malicious sites. Thus, spam filters became a requirement recently, important because the 1st line of take against the ever increasing drawback of unwanted email. Though varied techniques are developed, we tend to still receive them quite often because the spammers manipulate the filters. Also, the important mails might accidentally get caught within the filter. The tongue process is employed for the filtering of spam emails so as to boost on-line security. There are various approaches in human language technology to unravel this drawback. In one such approach, human language technology engine takes the unclassified email and its class as Associate in Nursing input then processes it more victimizing the applied

mathematics human language technology approach. Bayesian spam filtering is considered to be the most effective approach.

C. Text Classification:

Text classification is largely the classification of text based on the content of the text. Given associate input text, it predicts a pre-defined category label for it. A preferred example is spam filtering. Other examples include classifying the language of the supply text & genre of a fictional story.

Types of text classification:

- Binary classification: If there are unit precisely 2 categories to which a document belongs to, then it's a binary classification.
- Multi-class classification: If there are unit quite 2 classes & every document belongs to precisely one category, it is a multi-class classification.
- Multi-label classification: during this, a document has additional than one associated class within the classification scheme.

D. Speech Recognition:

A field of linguistics that involves developing methodologies & technologies that permits recognition of associate acoustic signals containing spoken language & interprets it into text. In straightforward words, it's the technique of changing spoken language into text. A language model is employed to make the text output that's conditioned on the audio information. When a person reads text into the system, the system analyzes the person's specific voice & uses it to fine-tune the popularity of that person's speech, therefore maintaining accuracy. Speech recognition is or else known as voice recognition. Advances in deep learning over the last ten years have allowed major players to deploy systems that involve:

- a) Transcribing a speech.
- b) making text captions for a show or T.V show.
- c) supply commands on radio whereas driving.
- d) Voice search.(Google, Siri)
- e) decision steering.
- f) automatic identification exploitation voice life science.

E. Sentiment Analysis:

Goal of sentiment analysis is to spot the sentiment among many posts or perhaps within the same post wherever feeling is not invariably expressly expressed. Sentiment analysis determines the angle, spirit, judgment or intent of the author. corporations use sentiment analysis to spot options

and sentiment on-line to know what their customers believe about their product and services and overall indicators of their name. It's done by distributing a polarity to the text – positive or negative – or making an attempt to acknowledge the underlying mood. Sentiment analysis understands sentiment in context to help USA higher perceive what's behind associate degree expressed opinion.

F. Document Summarization:

Information overload could be a real downside after we got to Access is a vital piece of knowledge, and already our access to data and knowledge so much exceeds our capacity to know it. Therefore a capability to summarize the meaning of documents and knowledge is changing into increasingly vital. Automatic summarisation is additionally relevant to grouping knowledge from social media associate degree additionally once accustomed to offer an overview of a point or diary post. Another desired outcome is to know deeper emotions. It reduces the redundancy from multiple sources.

G. Question Answering:

As speech understanding technology and voice input applications improve, the necessity for human language technology can solely increase. Search engines place the world's info at our fingertips, however still square measure primitive once involves truly answering specific queries announced by humans. Question responsiveness is changing into additional and additional standards, with the rising of applications like Siri, OK Google, chatbots and virtual assistants. Question responsive application is a system capable of coherently responsive a personality's request in language. it's going to be used as a text-only interface and as a spoken-dialogue system. Google has seen the frustration it's caused in users, who often have to strive for a variety of various search results to seek out the answer they're trying to find. Although definitely improving, this remains a relevant challenge for search engines, therefore changing into one amongst the most applications of NLP analysis.

H. Co-reference Resolution:

Co-reference resolution connects pronouns to the correct objects. It's essential to interpret the text properly. It is an important step for human language technology applications like document summarization, question responsiveness, etc. It happens once 2 or additional expressions during a text talk to the same person or factor. E.g. Bill aforesaid he would return, the right noun Bill and also the pronoun he refers to constant person,

i.e. Bill. Sense Disambiguation: Sense elucidation is that the downside of determining that sense (meaning) of a word is activated by the employment of the word during a specific context. While the human brain is pretty sensible at this task, a pc won't find it straightforward to acknowledge it. E.g. the pc won't be ready to acknowledge that the term pounds within the sentence I gained twenty pounds since the wedding! , presumably refers to the unit of mass instead of the currency.

I Named Entity Recognition (NER):

Named entity recognition could be a subtask of data extraction that locates and classifies named entities in text into predefined classes. These classes is also names of persons, organizations, locations, expressions of times, quantities, financial values, percentages, etc. Applying NER to a sentence is going to be ready to convert it from, Valium makes American state sleepyheaded, to makes the American state

J. Language Modeling:

A language model could be a performance that puts a chance measure over strings drawn from some vocabulary. It is a conditional distribution of crucial ith word in a very sequence of text, given all the previous word identities. Language modeling is employed in speech recognition, machine translation, part-of-speech tagging, data retrieval, parsing, etc.

MAJOR CHALLENGES IN NLP

The most natural means of communication between humans is the tongue, spoken, written ,or written. The dominance of the tongue as a way of communication among humans suggests that it'd be an associate degree agreeable medium in human-computer interaction. Thus, the major goal of human language technology would be the flexibility to use natural language as effectively as humans do. no tool can offer an associate degree in professional human quality word-sense disambiguation.

Also, the goal of human language technology is to alter computers to interact with themselves in communication and exploitation of natural human speech and language, so non-programmers will act with the computers simply and effectively.

When this goal is achieved, laptop systems are in a position to perceive, analyze, summarize, translate and generate accurate human text and language, which is the most natural means of communication between humans.

Human language technology considers the data structure of language rather than treating text as a sequence of symbols i.e., in human language technology many words build a phrase, many phrases build a sentence. So, to achieve this, goal computers should be blessed with natural language process capabilities, and these offer the following major challenges to the human language technology systems:

A. Machine Translation:

Machine Translation is the task of mechanically changing one linguistic communication into another linguistic communication, preserving the means of the input text and manufacturing fluent text within the output language. Correct translation requires not solely the flexibility to research and generate sentences in human languages but additionally human-like understanding of world information and context, despite the ambiguities of languages i.e., computers ought to be able to understand input in addition to one language, offer an output in more than one language and translate between languages. Words and phrases should be passed and taken in order that their meaning (as command, query, or assertion) may be determined Associate in Nursing an acceptable response is formulated and expressed. Application areas embrace science, diplomacy, transnational commerce and intelligence. Today, most IP resources ,and systems square measure offered solely for high resource languages (HRLs), like English, French and German. Whereas several low resource languages (LRLs) like Indonesian, Swahili-spoken and written by millions of individuals haven't any such resources or systems available. Thus a future major challenge for the IP community is to develop resources and tools for tons of ,and thousands of languages, not simply some.

B. Reading and Writing Text: Text reading and writing is one of the foremost challenges in NLP. Machine reading is the concept that machines may become intelligent, integrate and summarize info for humans, by reading ,and understanding the text offered i.e., computers ought to be able to perceive and method the data. Areas of applications embrace intelligence, logistics, office automation and libraries. With the emergence of the fashionable online world, we've got large storage of online info coded in human languages. E.g. scientific literature, where findings square measure still rumored nearly entirely in human language text. The standard of scientific literature is growing chop-

chops. So scientists are unable to stay up with the literature. Thus, we can say that the Associate in Nursing inflated the want for machine reading for the aim of comprehending and summarizing the literature in addition to extracting facts and hypotheses from this material. Also, machine reading has got to offer question-answering systems, by which humans will get answers from created information bases.

C. Interactive dialogue:

Since the Eighties, dialogue has been a preferred topic in IP research. Interactive dialogue permits humans straightforward, effective access to laptop systems, and victimization of linguistic communication for problem determination, deciding ,and management. Application areas embrace info access, command ,and management, factory control, workplace automation, provision ,and laptop power-assisted instructions. Human-machine interaction ought to be as natural, facile ,and multi-modal as interaction among humans. Early work on the text-based dialogue has currently been swollen to include spoken dialogue systems on mobile devices for information access and task-based applications (SDS). Although SDS work fairly well in restricted domains, where the topics of interaction square measure illustrious beforehand and wherever the words individuals square measure seemingly to use may be planned, they are not nonetheless terribly victorious in open domain interaction, where users could remark something in the slightest degree. The other challenges in building SDS square measure basic issues of recognizing and manufacturing traditional human colloquial behaviors.

D. Sentence Generation:

It is very non-trivial to come up with models that always generate grammatically correct and meaningful sentences.

LIMITATIONS OF NLP

In theory, information processing may be an engaging technique of human-computer interaction. Tongue recognition needs extensive information on the outside world.

Systems like SHRDLU that were developed earlier had restricted "block words" with restricted vocabularies and worked extraordinarily well. However, it had shortly failed once the systems were extended to additional realistic things with-real world ambiguity and complexity.

Hence, the development of information processing applications is difficult as computers need humans to "speak" to them in a very programming language that's precise, unambiguous

and highly structured.

The major issues with tongue process square measure ambiguity, unclerness and uncertainty.

These issues have to be compelled to be featured whether or not one is dealing with one sentence or discourse. The vital downside related to information processing is ambiguity. Ambiguity refers to associate expression (word/ phrase/ sentence) having over one interpretation. Ambiguity will occur

at the degree of lexical, syntactic, semantic, discourse ,and pragmatic analysis.

differing kinds of Ambiguity are:

1) Lexical Ambiguity:

It is the paradox of one word. A word will be ambiguous with relevance to its grammar category.

Ex: 1) She bagged a trophy.

2) She created a silver speech.

3) His worries had silvered his hair. The word 'silver' is employed as a noun, associate adjective, or verb.

2) Grammar ambiguity or Structural ambiguity: It is of 2 kinds: Scope ambiguity and Attachment ambiguity.

a) Scope ambiguity: Scope ambiguity involves operators and quantifiers.

Ex: recent men and girls were taken to safe locations. The scope of the adjective (i.e., the quantity of text it qualifies) is ambiguous. That is, whether or not the structure (old men and women) or ((old men) and women)? The scope of quantifiers is commonly not clear and it creates ambiguity. Ex: each man loves a lady. The interpretations will be, for each man there's a woman and additionally there is one explicit lady who is idolized by each man.

b) Attachment ambiguity: A sentence is claimed to own attachment ambiguity if a constituent fits over one position in a very breakdown tree. Attachment ambiguity arises from the uncertainty of attaching a phrase or clause to a part of a sentence.

Ex: the person saw the lady with the binoculars. It is ambiguous whether or not the person saw a lady carrying a binocular, or he saw her through his binocular. The meaning relies on whether or not the preposition 'with' is connected to the lady or the person.

3) linguistics Ambiguity:

This occurs once the means of the words themselves will be misinterpreted. There square measure 2 ways of reading the sentence, even when the syntax and therefore the meanings of the

individual words are resolved. Ex: Seema loves her mother and Sriya will too. The interpretations will be Sriya loves Seema's mother or Sriya likes her mother.

4) Pragmatic Ambiguity:

Pragmatic ambiguity happens once the context of a phrase gives it multiple interpretations .The problem involves the processing of extremely complicated tasks like user intention, sentiment, belief world, modals, etc.

Ex : I like you too.

This will be taken as

I like you (just such as you love me)

I like you (just like some other person does)

I like you (and i like somebody else)

I love you (as well as feel you).

III. FUTURE OF NLP

A. The bots:

Chatbots are ready to translate and interpret human language input. This is often done through a mixture of natural language processes and machine learning. There are well-known samples of AI and chatbots, e.g. Clever bot, Crotona and Siri. Chatbots are often employed in client services, chat bots facilitate customers get right to the purpose without the wait, respond to client queries ,and divert them to relevant resources. To be effective, chatbots should be quick, good and straightforward to use. To accomplish this, chat bots use NLP to grasp language typically over text or voice recognition interactions. The foremost obtrusive defect of today's chatbots is that they don't perceive what you're locution. They'll typically misinterpret what you sort, or ignore it fully. On the other hand, the chatbots that try to perceive and respond to every word you say will fail even more awkwardly. Thus, overcoming this issue of chat bots becomes a significant analysis area in NLP.

B. Support invisible UI:

An invisible UI (user interface) is wherever the user experience of a feature is thus intuitive that a UI isn't very required. a true- life example may be slippy doors, there is no assemblage needed, no button necessary ,and also the answer is invisible.

As packages and the web become additional helpful each day there is ever the additional ought to style invisible UIs. This level of intuition helps the user come through their goal faster and easier, then exploiting the user exploiting your application or website feeling glad. The idea of invisible UI depends on direct interaction

between the user and the machine. NLP leverages a bigger discourse understanding of human language. It gets higher understanding folks – despite how we say it and what we tend to be doing- which can be essential for any invisible UI application.

C. Smarter Search:

The future of NLP is additionally for smarter search. The same capabilities that permit a talk larva to grasp a customer's request will change the "search such as you talk" functionality.

IV. CONCLUSION

In our paper, we've given an outline of natural language processing and its applications. The impact of pc use of Natural Languages can have as profound a sway on society as would the breakthroughs in superconductors, cheap fusion or genetic engineering. The impact of natural language processing by machine is going to be larger than the impact of chip technology within the last twenty years, as linguistic communication is key to nearly all business, military & social activities. Therefore, the application of natural language processing has stopped.

REFERENCES

- Ms. Rijuka Pathak et al. Natural language processing approaches application and limitations. In Natural language processing approaches, application and limitations (Vol. 1 Issue 7, September – 2012)
- Julia Hirschberg et al. Advances in natural language processing. In SCIENCE (17 JULY 2015 • VOL 349 ISSUE 6245)
- Anjali M K1 et al. Ambiguities in Natural Language Processing. In International Journal of Innovative Research in Computer and Communication Engineering (Vol.2, Special Issue 5, October 2014)
- <http://www.expertsystem.com/natural-language-processingapplications/>
- <http://language.worldofcomputing.net/understanding/applications-of-natural-language-understanding.html>
- <http://www.ps.uni-saarland.de/~niehren/oz-natural-languagescript.html/vorlesung/node49.html>
- https://en.m.wikipedia.org/wiki/Natural_language_processing
- <https://blog.algorithmia.com/introduction-natural-languageprocessing-nlp/>
- Ralph Weischedel. White Paper on Natural Language Processing. In ResearchGate (November 2003)
- <http://www.expertsystem.com/natural-language-processingapplications/>
- <https://www.techemergence.com/natural-languageprocessing-business-applications/>
- <https://www.cl.cam.ac.uk/teaching/2002/NatLangProc/revised.pdf>
- <http://www.glossary.sil.org/term/pragmatics>
- [https://en.wikipedia.org/wiki/Morphology_\(linguistics\)](https://en.wikipedia.org/wiki/Morphology_(linguistics))
- <https://en.wikipedia.org/wiki/Phonology>
- <https://www.lifewire.com/applications-of-natural-languageprocessing-technology-2495544?print>
- https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_natural_language_processing.htm
- <http://language.worldofcomputing.net/machine-translation/machine-translation-process.html>
- https://en.m.wikipedia.org/wiki/Language_model
- https://en.m.wikipedia.org/wiki/Speech_recognition
- Geetika Sethi et al. Role of Natural Language Processing in Education. In IJSRE (Volume 4 Issue 3 March 2016)
- <https://machinelearningmastery.com/applications-of-deeplearning-for-natural-language-processing/>
- https://en.m.wikipedia.org/wiki/Speech_recognition
- <http://www.expertsystem.com/future-natural-languageprocessing/>
- Dr. Khaled M Alhawiti Natural Language Processing and its Use in Education. In International Journal of Advanced Computer Science and application (Vol. 5 number 12, 2014) [26] Ronan Collobert et al. Natural Language Processing (Almost) from Scratch. In Journal of Machine Learning Research
- Jill Burstein Opportunities of Natural Language Processing Research in Education. In Springer-Verlag Berlin Heidelberg 2009
- Dianne Litman Natural Language Processing for Enhancing Teaching and Learning. at AAAI Conference on Artificial Intelligence 2016