

TEXT AND SPEECH PROCESSING WITH NATURAL LANGUAGE PROCESSING

Prof. Nikhil V. Khandar

Assistant Professor, Dr Ambedkar Institute of Management Studies & Research College,
Deekshabhoomi, Bajaj Nagar, Nagpur, Maharashtra 440010

Email:nikhilrajputh@gmail.com

Prof. Amit B. Rehapade

Assistant Professor, Santaji Mahavidyalaya, Nagpur

Prof. Neha Mittal

Assistant Professor, Dr Ambedkar Institute of Management Studies & Research College,
Deekshabhoomi, Bajaj Nagar, Nagpur, Maharashtra 440010

ABSTRACT: -

Text analysis and natural language processing are crucial modules of artificial intelligence and various other types of technology in today's technological Ireland. Text mining, occasionally stated to as text analytics, converts amorphous text found in documents and records into normalised, organized facts that could be utilised for scrutiny or to sequence device learning procedures. Text mining utilises natural language processing (NLP). Growth of technology is accountable for supporting humans in obtaining favourable benefits from its usage. This essay convicts' artificial intelligence of using text analysis and natural language processing. Natural language a sort of semantic that aids in text reading by encouraging human comprehension of a language that is naturally spoken, such as English or another type. When conducting research, the existing literature and studies were taken into account in order to obtain a deeper understanding of the subject using the content analysis method. The study's findings have convinced researchers that text analysis and natural language processing are extra crucial than previously thought for artificial intelligence.

Keywords: Natural Language Processing, Artificial intelligence, Text analysis, etc.

INTRODUCTION

In its simplest form, natural language processing is the area of artificial intelligence that helps processors comprehend, manipulate, and interpret human language. Natural language treating brings together a varied series of academic fields, counting processor science and computational linguistics, in an effort to close the communication gap between humans and computers. The method of extracting information from a bigger amount of textual material is simply described as artificial intelligence text analysis. Semantic analysis and syntactic analysis are the two main categories of NLP approaches. Using fundamental grammatical concepts, syntactic analysis, also known as parsing, examines text to find sentence structures, organise words, and determine how they link. Text apprehension is necessary for semantic examination. It starts by analysing the importance of every expression (lexical semantics). When it comes to automating the entire process of comprehending client remarks on a broader scale, natural language processing is the possibility that holds the most promise for organisations. It will assist them in making decisions based on facts in order to improve the company.

The significance of NLP can be understood in terms of how it enables computers to converse with people in their individual vernaculars and scales various kinds of jobs in respect to vernaculars. As an illustration, the definition of natural language processing is that it enables all computers to read text, hear speech, analyse it, measure sentiment, and identify the sections of the text that need to be occupied of. Human languages are incredibly diverse and multifaceted. We have been expressing ourselves in a wide variety of vocal and written means. The resolution of linguistic ambiguity and the

addition of a numerical structure that is considerably more beneficial for the records to arguments in many types of downstream presentations, such as speech appreciation or transcript analytics, make natural language processing particularly significant in artificial intelligence.

LITERATURE REVIEW

(Kongthon,2009) suggested use of artificial intelligence's natural language processing in the creation of an online tax system. This specific application was used to demonstrate the viability of using text investigation and natural language processing in artificial intelligence to safeguard the upcoming. The widely held of high-level applications for natural language processing integrate elements that mimic smart behaviour, such as the deceptive understanding of natural languages.

After reviewing existing approaches in this area, Jean (2014) offers a solution based on significance sampling that enables us to practice a big terminology lacking collective NMT model's training difficulty. This method can be used to resolve machine transformation problems. Then, via a single neural system that has been optimised for conversion presentation, they advise an estimate exercise approach built on (biased) sampling that will agree you to train an NMT model with a far broader target wording.

A applied approach to natural language processing (NLP) may be challenging to develop, create, or even approach. According to (Collobert, 2011), their contribution to this field of study represents a crucial turning point in the discussion of lined training algorithms that greatest take use of the tremendous developments in computer hardware. These writers in this case are constructed on sizable, unlabelled fact sets using the NLP tagger, and the training algorithm finds inside portrayals that are beneficial for entire professions.

(Vinyals, 2015) Use a recurring neural network with a consideration 4 method to create phrase analyse trees to appropriately perform extremely focused study on foreign language grammar. A fundamental problem with a wide range of submissions in linguistics and natural language processing is syntactic constituency parsing. Since this problem has been the topic of extensive research for several years, there are now exceptionally accurate domain-specific parsers available.

(Fisher, 2010) Text mining is the procedure of looking through sizable amount or group of written means to create novel information. The goal of text mining, for instance, is to find significant info in text by altering text into information that might be used for more research. One technique used in text mining to accomplish this is NLP. Any programme for natural language processing must have a thesaurus, lexicon, ontology, and present articles.

(Mnasri, 2019) One of the most popular uses of text mining and NLP is social media observing, which involves examining collection of user-generated content to find sentiment, feelings, and consciousness about a particular subject.

(Moreno,2016) A gathering of methods for dispensation such data to decipher fundamental meaning is known as natural language processing (NLP). Applications that categorise, excerpt structure, summarise, and interpret data—which can be express, text, or even an image—are created using machine learning (ML) techniques.

METHODOLOGY

Objectives of the study

The following were few of the main objectives taken into account during the research:

- To investigate Text Analytics and the methods used in AI;
- To study NLP in AI and the many techniques employed.

Research Methodology

The content analysis approach is recognised for doing the analysis of already published material and for having in-depth accepting of Text Analytics and Natural Language Processing in artificial intelligence. The literature that was already in existence has been removed with the aid of keywords after taking into consideration various research papers and earlier analyses. Information from those works was extracted, and when it was taken out, it was confirmed for exactness and dependability by means of the right foundations. Specifics were included in the paper and immaterial information was eliminated after accuracy test. Subsequently, the addition of the applicable material in the article, succeeding scrutiny has been attained concerning this specific issue.

NLP AND TEXT ANALYTICS

Text analysis, specifically including, grouping, and categorising words to extract the construction and significant content from greater amounts of content, is defined as working hand in hand with natural language processing. Text analysis is used to examine the text's content and originate new kinds of variables from unprocessed facts which can then be visualised, clarified, and utilised as ideas for forecasting prototypes and other kinds of arithmetical techniques. Text analysis and natural language processing are combined in a wide range of applications, including:

- Investigation discovery: locating patterns and clues in emails or reports that were made to aid in the process of identifying and resolving crimes that had occurred.
- Subject matter expertise: the organisation of the content into relevant categories so that readers will be able to act and identify patterns.
- Social media analytics: observing sentiment and awareness of particular issues, as well as the identification of major influencers.

For successful readers, the ability to analyse the text is crucial. The process of analysing a text entails dissecting its themes and organisational principles in order to better comprehend, evaluate, and draw conclusions from it. It is evident that text analysis in artificial intelligence moving in the course of dissecting the concepts and organisation of the text so that a person using it can know it clearly.

Natural Language Processing

Computers can now understand human language thanks to a process called natural language processing (NLP). Behind the scenes, NLP analyses the traits and importance of the sentences, using algorithms to glean meanings and deliver information. In other words, numerous tasks can be carried out automatically using human language.

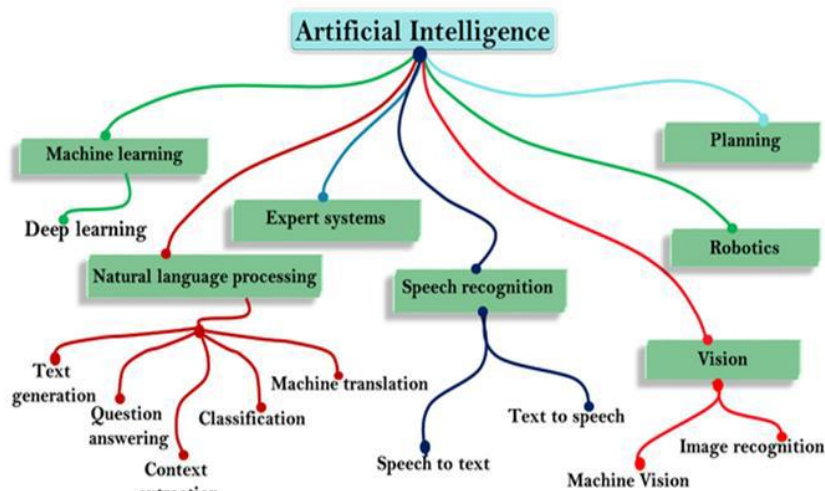


Fig1: NLP in AI

The study of natural language processing (NLP) focuses on how computers and translators comprehend social language. With NLP, computers can comprehend spoken or written language and carry out responsibilities like translation, keyword extraction, subject categorization, and more.

The amount of pertinent information in the questions has an impact on how accurate the response is.

- Users can ask questions on any subject and receive direct responses in a matter of seconds.
 - Deployment is easy.
- Using software is less expensive than hiring a person.
- The NLP system responds to natural language queries, which a human may take two or three times longer to complete than a machine.
- Enable you to process more language-based data impartially and consistently (as opposed to a person) without getting tired.

The NLP technique measures other language processes and enables a computer to converse in a person's language.

NLP Techniques

In order to support processors, figured out text Natural Language Processing (NLP) usages two techniques: syntactic analytics as well as semantic scrutiny.

Syntactic Analysis

Syntactic analysis - or parsing - inspects text by means of essential linguistic values to perceive a judgement construction, organize arguments and how they attach.

Some of its main subtasks are:

- Tokenization includes separating a text up into lesser pieces called tokens (which may be phrases or words) to simplify handling of material.
- Control the meaning of words with the help of part of speech tag labels such as substance, verb, adjective, adverb, etc.(e.g- the term "book" denotes to numerous objects whether working as a verb or a substantive).
- Lemmatization & stemming contains in reducing varied phrases to their basic form to ease analysis.
- Stop-word removal often removes words that don't contribute semitone value, such as I, they, have, etc.

Semantic analysis

Semantic analysis requires text detention. It begins by evaluating each phrase's significance (lexical semantics). The context is then examined to determine the arrangement and meaning of the words. These are the primary responsibilities of semantic analysis. In order to determine what a word means in a certain context, word meaning disambiguation is used.

Applied NLP

The following are some of the main applications of NLP:

• Language Translation

Machine translation technology has evolved significantly in recent years, with Facebook's translations performing at a level of superhuman performance in 2019. Companies can engage in a variety of languages, improve their global communication, or access new markets thanks to translation tools.

• Text extraction:

You can extract predefined text information using text extraction. When working with enormous amounts of data, this programme aids in the identification and extraction of keywords, significant attributes (such as product codes, colours, and specs), and called entities.

- Chatbot

Chatbots are artificial intelligence programmes that can speak or text with people. Because they can handle many inquiries at once, are available 24/7 (which reduces response times), and can free up human staff members from continually responding to the same questions, chatbots are increasingly used in customer support. Since chatbots learn from each encounter and expand operator determined, you can rely on them to execute routine and easy tasks.

Text analytics

Automated text analysis is text analysis. There is a difference amongst text analysis and text mining, despite the two terms occasionally being used interchangeably. While you wait, text analysis delivers you with more in-depth measurable data so you can make wise conclusions. Text analysis delivers awareness into movements, patterns, and consumer attitudes when applied to formless evidence, also known as open-ended input, to find and prioritise ways to progress customer knowledge

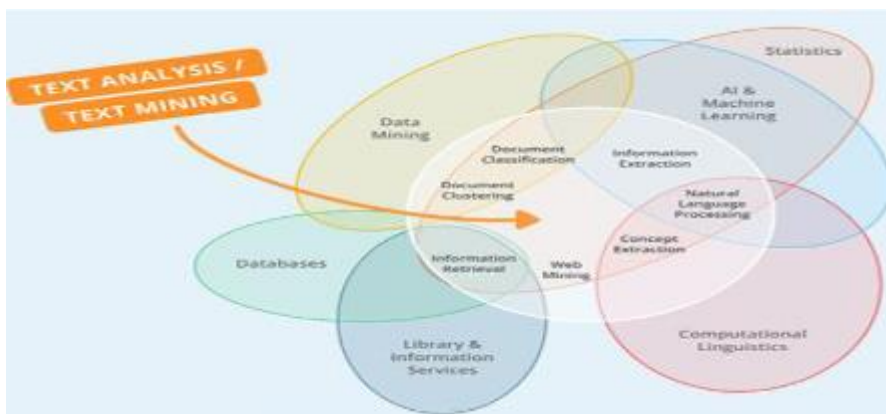


Fig2: Domain and Subdomain of Text Analysis

Unstructured feedback has the advantage of being consistent with what your customers have to say and providing the most insight into what they are thinking. There are numerous ways in this response for you to absorb as much information as you can.

Text analysis methods

Just a few of the key technologies at play are listed below:

Artificial intelligence (AI) is the ability of a computer system to bring out jobs that typically require human brain. The list of these duties does not just stop at decision-making and voice recognition. To make it easier to review your unstructured remarks, Text Analytics automatically classifies and categorises massive amounts of text.

- Computer learning (ML)

Even while machine learning is a part of AI, it is separate. The main goal of machine learning (ML) is to develop algorithms for computers that can efficiently study from experience and automatically adjust to improve performance without human programming. Text analytics uses machine learning to determine in what way to classify new text pieces based on before treated text and to determine whether the categories used to classify these text pieces need to be changed in light of the patterns found.

- *Deep learning (DL)*

A regulated, focused subclass of artificial intelligence that includes a computer system's ability to analyse data and draw conclusions about other data. Deep learning may be used in text analytics to

improve the precision of autonomous text analysis and better understand the context in unstructured feedback.

Sentiment Evaluation

Using text analysis and NLP to process messages and automatically determine if they have a good, negative, or neutral attitude. Sentiment analysis is a great place to start when assessing your disorganised feedback because it helps you quickly identify important and developing issues as well as areas where your interviewers may improve.

• *Classification of text based on rules*

Using a lexicon or word dictionary to instruct a computer system on how to categorise new resources Text analytics uses rule-based text categorization to automatically assign emotions or subjects to every text it analyses.

CONCLUSION

According to research done, it can be conclusively proven that NLP is a far better way than other methods since it can recognize both text and speech, whereas other methods, such as text mining, only evaluate text quality. While knowledge of multiple topics like cosine similarity or feature hashing, and text processing, like Perl or Python, is necessary for text mining, it is less necessary for NLP systems to have abilities like NLTK or competency in neural networks. Understanding statistical methods is another crucial component of text mining. Conferring to research shown over the previous ten years, it is evident that NLP is used more often than other geometric methods since it is more practical, user-friendly, and requires less understanding.

REFERENCES

- Ahonen, Helena, OskariHeinonen, Mika Klemettinen, and A. InkeriVerkamo. Applying data mining techniques in text analysis. Report C-1997-23, Dept. of Computer Science, University of Helsinki, 1997.
- Cavazza, Marc, SrikanthBandi, and Ian Palmer. "Situated AI in video games: integrating NLP, path planning and 3D animation." In AAAI 1999 Spring Symposium on Artificial Intelligence and Computer Games, pp. 6-12. 1999.
- Nikhil Khandar and Jitendra Sheethlani "Sentiment Analysis Approach In Natural Language Processing For Data Extraction" in International Journal of Computer Engineering and Technology 13(2), 2022, pp. 27-32. <https://iaeme.com/Home/issue/IJCET?Volume=13&Issue=2, 13-July-2022>.
- Day, M. Y. (2020). Artificial Intelligence for Text Analytics.
- Fisher, Ingrid E., Margaret R. Garnsey, SunitaGoel, and Kinsun Tam. "The role of text analytics and information retrieval in the accounting domain." Journal of Emerging Technologies in Accounting 7, no. 1 (2010): 1-24.
- Ise, Orob Anderson. "Integration and analysis of unstructured data for decision making: Text analytics approach." International Journal of Open Information Technologies 4, no. 10 (2016).
- Lalwani, T., Bhalotia, S., Pal, A., Rathod, V., &Bisen, S. (2018). Implementation of a Chatbot System using AI and NLP. International Journal of Innovative Research in Computer Science & Technology (IJIRCST) Volume-6, Issue-3.
- Lehnert, Wendy, and Beth Sundheim. "A performance evaluation of text-analysis technologies." AI magazine 12, no. 3 (1991): 81-81.
- Lee, L. (2002, July). A non-programming introduction to computer science via NLP, IR, and AI. In Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics (pp. 33-38).
- Mnasri, Maali. "Recent advances in conversational NLP: Towards the standardization of Chatbot building." arXiv preprint arXiv:1903.09025 (2019).

- Mamou, J., Pereg, O., Wasserblat, M., Eirew, A., Green, Y., Guskin, S., ... & Korat, D. (2018). Term set expansion based nlp architect by intel ai lab. arXiv preprint arXiv:1808.08953.
- Moreno, A., & Redondo, T. (2016). Text analytics: the convergence of big data and artificial intelligence. *IJIMAI*, 3(6), 57-64.
- Yang, Yiwei, EserKandogan, Yunyao Li, Prithviraj Sen, and Walter S. Lasecki. "A study on interaction in human-in-the-loop machine learning for text analytics." In *IUI Workshops*. 2019.