International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET2310214

A Review Study on Applications of Natural Language Processing

Abhay A. Dande[1], Dr. M. A. Pund[#2]

[1]Ph.D Research Scholar, Computer Science & Engineering, PRMIT&R- Badnera, Maharashtra, India

[2]Professor, Computer Science & Engineering, PRMIT&R- Badnera, Maharashtra, India

ARTICLEINFO

Article History:

Accepted: 05 March 2023
Published: 20 March 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
122-126

ABSTRACT

Natural Language Processing (NLP) is a subfield of computer science and artificial intelligence (AI) that deals with the interaction between computers and humans in natural language. The goal of NLP is to enable computers to understand, interpret, and generate human language, allowing them to perform tasks that would typically require human intelligence. In recent years, NLP has experienced significant growth due to the increasing availability of digital content and the need for intelligent information extraction from these sources. This paper reviews recent advances in NLP, including text classification, named entity recognition, sentiment analysis, machine translation, and speech recognition. Additionally, it highlights some of the challenges and future directions for NLP research.

Keywords: Named Entity Recognition, Word Embedding, Chabot, Machine Translation

I. INTRODUCTION

Language is a complex system, and NLP involves a range of techniques, including machine learning, deep learning, and linguistics, to analyze and understand it. The primary objective of NLP is to build machines that can process natural language input and output in a way that is both accurate and effective.

Natural Language Processing (NLP) is a subfield of artificial intelligence that deals with the interaction between humans and computers using natural language. The goal of NLP is to develop algorithms that can understand, analyze, and generate human language. In recent years, NLP has gained a lot of attention due to the increasing use of digital content and the need for intelligent information extraction from these sources.

The goal of NLP is to enable humans to communicate with machines in a more intuitive and efficient way. In recent years, NLP has gained significant attention due to the widespread use of digital content and the need for intelligent information extraction from these sources

NLP has a broad range of applications, including language translation, chatbots, sentiment analysis, voice recognition, and more. NLP has also been used in fields such as healthcare, finance, and law, where natural language data is abundant and valuable. Overall, NLP is a rapidly growing field with many exciting possibilities for the future. As natural language

data continues to proliferate, NLP is becoming increasingly important in a wide range of industries and applications, and it is likely that the demand for NLP expertise will only continue to grow in the years to come.

Some of the key challenges in NLP include the ambiguity of natural language, the variability of language use, and the need to consider context and meaning when analyzing language. Researchers in NLP are constantly working to develop new techniques and models to overcome these challenges and improve the accuracy and effectiveness of NLP systems.

II. NLP Applications Use Cases

1. Text Classification

Text classification is a fundamental task in NLP, which involves categorizing a piece of text into predefined categories based on its content. The categories can include topics, sentiments, and other attributes. Various machine learning algorithms have been proposed to solve this task, including Naive Bayes, Support Vector Machines, and Neural Networks. In recent years, deep learning approaches, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown significant improvements in text classification tasks. This algorithm categorizes a text into one or more predefined categories or topics, based on its content.

Text classification has many applications in various domains, including sentiment analysis, spam detection, and news classification. Various machine learning algorithms have been proposed to solve this task, including Naive Bayes, Support Vector Machines, and Neural Networks.

2. Named Entity Recognition

Named Entity Recognition (NER) involves identifying and classifying named entities such as people, organizations, and locations in text. This task is essential in various applications, such as information extraction, text summarization, and question answering. Different machine learning algorithms have been proposed for NER, including Conditional Random Fields, Hidden Markov Models, and Maximum Entropy models. Recently, deep learning approaches, such as LSTM and BiLSTM, have shown promising results in NER tasks.

Named Entity Recognition (NER) is the task of identifying and classifying named entities such as people, organizations, and locations in text. It is an essential task in various NLP applications, such as information extraction, text summarization, and question answering. In recent years, deep learning approaches have shown significant improvements in NER. For example, the BiLSTM-CRF model proposed by Huang et al. in 2015 combines Bidirectional Long Short-Term Memory (BiLSTM) with Conditional Random Fields (CRF) to achieve state-of-the-art performance on NER

Different machine learning algorithms have been proposed for NER, including Conditional Random Fields, Hidden Markov Models, and Maximum Entropy models. Named Entity Recognition (NER) - This algorithm identifies named entities in a given text, such as people, organizations, locations, and dates.

3. Sentiment Analysis

Sentiment analysis is a task in NLP that involves identifying and categorizing the opinions and emotions expressed in text. The categories can include positive, negative, and neutral sentiments. Sentiment analysis has various applications, including brand reputation management, product marketing, and political analysis. Various machine learning algorithms have been proposed for sentiment analysis, including Naive Bayes, Support Vector Machines, and Neural Networks. In recent years, deep learning approaches, such as CNNs and RNNs, have shown significant improvements in sentiment analysis tasks.

Sentiment analysis is a popular NLP task that involves determining the sentiment expressed in a piece of text. It has many applications in various domains, such as social media analysis and product reviews. One of the most influential works in this field is the VADER

(Valence Aware Dictionary and sEntiment Reasoner) algorithm proposed by Hutto and Gilbert in 2014. VADER is a rule-based algorithm that uses a lexicon-based approach to sentiment analysis. This algorithm determines the overall tone or sentiment of a text, whether it is positive, negative, or neutral.

4. Machine Translation

Machine Translation is the task of translating text from one language to another automatically. This task has become increasingly important due to the global nature of business, science, and culture. Various machine learning algorithms have been proposed for machine translation, including Statistical Machine Translation and Neural Machine Translation. Recently, Transformer-based models, such as BERT and GPT-2, have shown significant improvements in machine translation tasks.

Machine translation is the task of translating text from one language to another using computer algorithms. It is a challenging task in NLP, as it requires understanding the semantic meaning of the text in both languages. In recent years, neural machine translation (NMT) has shown significant improvements in this task. The attention mechanism proposed by Bahdanau et al. in 2015 is a crucial component of NMT that allows the model to focus on specific parts of the source sentence when generating the target sentence.

5. Speech Recognition

Speech Recognition is the task of recognizing and transcribing spoken language into text. This task has many applications, including virtual assistants, transcription, and language learning. Various machine learning algorithms have been proposed for speech recognition, including Hidden Markov Models, Neural Networks, and Convolutional Neural Networks. Recently, deep learning approaches, such as LSTM and Transformer-based models, have shown significant improvements in speech recognition tasks.

6. Word Embedding

Word embedding are a popular technique used in NLP for text representation. The idea behind word embedding is to represent each word as a vector in a high-dimensional space, where the position of the vector captures the semantic meaning of the word. One of the most influential works in this field is the Word2Vec algorithm proposed by Mikolov et al. in 2013. Word2Vec is a neural network-based algorithm that learns word embedding from a large corpus of text data. Since its introduction, Word2Vec has been used in various NLP tasks, such as sentiment analysis, question answering, and machine translation.

7. Context-based Applications

Context-based NLP algorithms refer to natural language processing algorithms that take into account the context of the text being analyzed. These algorithms are designed to identify patterns and relationships between words and phrases within a given context, enabling more accurate and relevant analysis. Context-based NLP algorithms can be used for a variety of tasks, such as sentiment analysis, text classification, and language translation. For example, sentiment analysis algorithms can identify the sentiment of a text by analyzing the words and phrases used within a given context. Text classification algorithms can identify the topic of a text by analyzing the key words and phrases used within a given context. Context-based NLP algorithms use various techniques, such as machine learning and deep learning, to analyze text data. These techniques enable the algorithms to learn from past data and improve their accuracy over time. Overall, context-based NLP algorithms are an essential tool for businesses and organizations looking to extract valuable insights from text data. By taking into account the context of the text being analyzed, these algorithms enable more accurate and relevant analysis, leading to more informed decision-making

8. Part of Speech Tagging

This algorithm tags each word in a text with its part of speech, such as noun, verb, adjective, or adverb. It has very high importance in text analysis and the advance concept of applications such as text summarization and topic modelling. It plays a vital role in document and text parsing. It makes the use of the concepts of

stemming and lemmatization from the python nltk library for this purpose. More advanced libraries from deep learning are now also helpful in this process to for effective and efficient tagging.

9. Topic Modeling

This algorithm identifies the main topics and themes in a given text, based on the frequency and co-occurrence of certain words and phrases. It has the dictionary based and vocabulary based approaches in which the list of the topics can be made available. Depending upon the context and frequency of the words they can be mapped to the specific mentioned topic and can be easily displayed using some of the concepts of machine learning such as wordcloud.

10. Word Sense Disambiguation

This algorithm resolves the ambiguity of a word with multiple meanings, by identifying the correct sense of the word in a given context. It makes the use of Long Short Term Memory (LSTM), Bidirectional LSTM and deep learning concepts for this purpose and are able to preserve the context of the word for large number of related sentences and paragraphs.

In recent years, deep learning approaches have shown significant improvements in NLP tasks. Deep learning involves training neural networks with multiple layers to learn hierarchical representations of text data. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are two popular deep learning architectures used in NLP. CNNs are commonly used in tasks such as text classification and sentiment analysis, while RNNs are used in tasks such as language modeling, machine translation, and speech recognition.

III. Conclusion

In this review paper, we discussed some of the recent advancements in NLP and highlighted some of the most influential works in the field. Word embedding, named entity recognition, sentiment analysis, and machine translation are some of the most important tasks in NLP. Deep learning approaches have shown significant improvements in various NLP tasks, and further research in this field is needed to address the remaining challenges. Despite the progress made in NLP, there are still several challenges in this field. One of the significant challenges is the lack of labeled data for many languages and domains. This challenge has led to the development of unsupervised and semi-supervised learning approaches in NLP, which attempt to use unlabeled data to learn useful representations of text data. Another challenge is the interpretability of deep learning models, which makes it difficult to understand how these models make predictions.

IV. Future Scope

NLP has many applications in various domains, and there have been significant advancements in this field in recent years. Machine learning algorithms and deep learning approaches have been proposed to solve various NLP tasks, including text classification, named entity recognition, and sentiment analysis. However, there are still several challenges in NLP, such as the lack of labeled data and the interpretability of deep learning models. Further research in this field is needed to address these challenges and advance the state of the art in NLP.

V. REFERENCES

[1]. E.D. Liddy, Natural Language Processing, 2001.

[2]. N. Kaur1, V. Pushe and R Kaur,"Natural Language Processing Interface for Synonym", International Journal of Computer Science and Mobile Computing, Vol.3 Issue.7, July- 2014, pp. 638-642 ,ISSN 2320–088X.

[3]. E. Hovy, J.Lin, J.Prager, D. Radev, L.Vanderwende, R.Weischedel, "Natural Language Processing", This report is one of five reports that were based on the MINDS workshops.

[4]. G.Chowdhury, "Natural language processing", Annual Review of Information Science and

Technology, 2003, 37. pp. 51-89, ISSN 0066-4200.

[5]. P. Jackson and I. Moulinier,"Natural Language Processing for Online Applications": Cambridge University press, New York.2012, page 7-9.

[6]. R. Bose. "Natural language processing: Current state and future directions". International Journal of the Computer, the Internet and Management Vol. 12#1 (January – April, 2004) pp. 1 – 11

[7]. K. W Church and L.F Rau," Commercial applications of Natural Language Processing". Communication of the ACM, vol 38, No. 11,November 1995

[8]. R. J Bobrow and R.M. Weischedel, "Challenges in Natural Language Processing," Cambridge University press, New York.1993

[9]. R. Mihalcea, H. Liu, and H. Lieberman, "NLP (Natural Language Processing) for NLP (Natural Language Programming), " In Proceedings of CICLING'06, LNCS, Springer, 2006, pp. 319-330.

[10]. P. M Nadkarni, L. O. Machado and W. W. Chapman, "Natural Language Processing: An Introduction,"J Am Med Inform Assoc, vol.18, 2011, pp. 544-551.

[11]. B. Manaris, "Natural Language Processing: A Human–Computer Interaction Perspective," Appears in Advances in Computers (Marvin V. Zelkowitz, ed.), Academic Press, New York, vol. 47, 1998, pp. 1-66.

[12]. C. Wu and Y. Chen, "A Survey of Researches on the Application of Natural Language Processing in Internet Public Opinion Monitor," Proceedings of International Conference on Computer Science and Service System (CSSS), IEEE, 2011, pp. 1035-1038

Cite this article as :

Abhay A. Dande, Dr. M. A. Pund#, "A Review Study on Applications of Natural Language Processing ", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 10 Issue 2, pp. 122-126, March-April 2023. Available at doi : https://doi.org/10.32628/IJSRSET2310214
Journal URL : https://ijsrset.com/IJSRSET2310214