Journal of Nonlinear Analysis and Optimization Vol. 14, Issue. 2, No. 2: 2023 ISSN : **1906-9685**



IMPORTANT ROLE OF MACHINE LEARNING IN NATURAL LANGUAGE PROCESSING

Ms.T.Nithya M.Sc., B.Ed, Assistant Professor, PG Department of Computer Applications, Marudhar Kesari Jain College for Women, Vaniyambadi.

Abstract:

Natural Language Processing (NLP) is the part of Artificial Intelligence that helps computers communicate with humans in their own language and scales other language-related tasks. NLP makes it possible for computers to read text, hear speech, interpret it, measure sentiment, and determine which parts are important. Combined with machine learning algorithms, NLP generates systems that learn to perform tasks on their own and get better through experience. NLP-powered tools can help you to classify social media posts by sentiment, or extract named entities from business emails, among many other things. The role of machine learning in the NLP is clearly defined in this paper with its application areas.

Keywords : Security, Encryption, Block chain

Introduction

Machine Learning and Natural Language Processing are important subfields of Artificial Intelligence that have gained prominence in recent times. The goal of NLP is to build systems that can make sense of the text and automatically perform tasks like translation, spell check, or topic classification. Machine Learning and Natural Language Processing play an awfully important part in making a synthetic agent into an artificial 'intelligent' agent. An Artificially Intelligent system can accept better information from the environment and might act on the environment in a user-friendly manner because of the advancement of Language Processing. Machine Learning gives the system the power to find out past experiences and examples. General algorithms perform a hard and fast set of executions in line with what it's been programmed to try, and they don't possess the flexibility to resolve unknown problems. And, in the real world, most of the issues faced contain many unknown variables, making the standard algorithms very less effective. This is often where machine learning comes to the fore. With the assistance of past examples, a machine learning algorithm is much better equipped to handle such unknown problems.

Differences between Natural Language Processing and Machine Learning

Although Natural Language Processing, <u>Machine Learning</u>, and Artificial Intelligence are sometimes used interchangeably, they have different definitions. AI is an umbrella term for machines that can simulate human intelligence, while NLP and ML are both subsets of AI.

Artificial Intelligence: is a part of the greater field of Computer Science that enables computers to solve problems previously handled by biological systems. AI has many applications in today's society. NLP and ML are both parts of AI.

Natural Language Processing: is a form of AI that gives machines the ability to not just read, but to understand and interpret human language. With NLP, machines can make sense of written or

spoken text and perform tasks including speech recognition, sentiment analysis, and automatic text summarization.

<u>Machine Learning</u> : is an application of AI that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine Learning can be used to help solve AI problems and to improve NLP by automating processes and delivering accurate responses.



How Natural Language Processing can be applied

Some NLP-based solutions include translation, speech recognition, sentiment analysis, question/answer systems, chatbots, automatic text summarization, market intelligence, automatic text classification, and automatic grammar checking. These technologies help organizations to analyze data, discover insights, automate time-consuming processes and/or gain competitive advantages.

➤ Translation

Translating languages is more complex than a simple word-to-word replacement method. Since each language has grammar rules, the challenge of translating a text is to do so without changing its meaning and style. Since computers do not understand grammar, they need a process in which they can deconstruct a sentence, then reconstruct it in another language in a way that makes sense. Google Translate is one of the most well-known online translation tools. Google Translate once used Phrase-Based Machine Translation (PBMT), which looks for similar phrases between different languages. At present, Google uses Google Neural Machine Translation (GNMT) instead, which uses ML with NLP to look for patterns in languages.

Speech recognition

Speech recognition is a machine's ability to identify and interpret phrases and words from spoken language and convert them into a machine-readable format. It uses NLP to allow computers to simulate human interaction, and <u>ML</u> to respond in a way that mimics human responses. Google Now, Alexa, and Siri are some of the most popular examples of speech recognition. Simply by saying 'call Jane', a mobile device recognizes what that command means and will now make a call to the contact saved as Jane.

Sentiment Analysis

Sentiment analysis uses NLP and <u>ML</u> to interpret and analyze emotions in subjective data like news articles and tweets. Positive, negative, and neutral opinions can be identified to determine a customer's sentiment towards a brand, product, or service. Sentiment analysis is used to gauge public opinion, monitor brand reputation, and better understand customer experiences.

The stock market is a sensitive field that can be heavily influenced by human emotion. Negative sentiment can lead stock prices to drop, while positive sentiment may trigger people to buy more of the company's stock, causing stock prices to increase.

Chatbots

Chatbots are programs used to provide automated answers to common customer queries. They have pattern recognition systems with heuristic responses, which are used to hold conversations with humans. Initially, chatbots were used to answer basic questions to alleviate heavy volume call centers and offer quick customer support services. But AI-powered chatbots are designed to handle more complicated requests making conversational experiences increasingly intuitive. Chatbots in healthcare, for example, can collect intake data, help

patients assess their symptoms, and determine next steps. These chatbots can set up appointments with the right doctor and even recommend treatments.

Question-Answer Systems

Question-answer systems are intelligent systems that are used to provide answers to customer queries. Other than chatbots, question-answer systems have a huge array of knowledge and good language understanding rather than canned answers. They can answer questions like "When was Abraham Lincoln assassinated?" or "How do I get to the airport?" and can be created to deal with textual data, audio, images and videos.

Question-answer systems can be found in social media chats and tools such as Siri and IBM's Watson. In 2011, IBM' Watson computer competed on Jeopardy, a game show during which answers are given first, and the contestants supply the questions. The computer competed against the show's two biggest all-time champions and astounded the tech industry when it won first place.

Automatic text summarization

Automatic text summarization is the task of condensing a piece of text to a shorter version, by extracting its main ideas and preserving the meaning of content. This application of NLP is used in news headlines, result snippets in web search, and bulletins of market reports.

Market Intelligence

Market Intelligence is the gathering of valuable insights surrounding trends, consumers, products and competitors to extract actionable information that can be used for strategic decision-making. Market Intelligence can analyze topics, sentiment, keywords, and intent in unstructured data and is less time consuming than traditional desk research. Using Market Intelligence, organizations can pick up on search queries and add contextually relevant synonyms to search results. It can also help organizations decide which products or services to discontinue or what customers to target.

Automatic text classification

Automatic text classification is another fundamental solution of NLP. It is the process of assigning tags to text according to its content and semantics which allows for rapid, easy retrieval of information in the search phase. This NLP application can differentiate spam from non-spam based on its content.

Automatic grammar checking

Automatic grammar checking, the task of detecting and correcting grammatical errors and spelling mistakes in text depending on context, is another major part of NLP. Automatic Grammar Checking will alert you to a possible error by underlining the word in red.

Methodology for Machine Learning

Machine learning (ML), or deep learning, depends on algorithms to inform what actions are taken and then produce an inferred function. In the future, we may see machines achieve true self-awareness and operate independently from human, data-influenced input. But for now, humans and data will continue to play a critical role in shaping machine-driven predictions. There are two main methods to guide your machine learning model—supervised and unsupervised learning. Depending on what data is available and what question is asked, the algorithm will be trained to generate an outcome using one of these methods. The difference between them is that supervised learning uses a full set of labeled data during training. In unsupervised learning, the data set is provided without explicit instructions on what to do with it; the machine is basically winging it.

Supervised learning

The supervised learning technique is more commonly used in <u>machine learning</u> because it deals with straightforward tasks and is easy to implement. Data inputs are labeled with the answer that the algorithm should arrive at, which helps the machine pick out patterns in the future, better differentiate data, or make predictions. Supervised learning is classified into two categories of algorithms and is ideal for problems where there are reference points available.

1. Classification: A classification problem exists when the output variable is a specific category.

20

21

Vol. 14, Issue. 2, No. 2: 2023

- 2. **Regression:** A regression problem exists when the output variable is a real value that fluctuates (i.e. dollars, weight, measurement).
- Unsupervised learning

In this technique, the machine learning model learns organically instead of receiving a data set with explicit instructions. The unsupervised learning model is classified into four different categories of algorithms, which group data based on similarities or relationships among variables:

- 1. **Clustering:** The deep learning model looks for data and features that are similar then groups them together.
- 2. Association: By reviewing key attributes in the data, an unsupervised learning model can predict other attributes that they're commonly associated with.
- 3. Anomaly detection: In this instance, the model is used to call attention to data outliers. For instance, banks detect fraud by looking for unusual purchase patterns with customers—like if a card is used in two very different locations in one day, the bank notices and investigates the activity.
- 4. Artificial neural networks (or autoencoders): An autoencoder takes input data, compresses it into code, and then tries to recreate the input from that code while removing any signal noise so data quality is improved.

Less-known machine learning methods

There are other, less common methods for machine learning that we're starting to see used more frequently, perhaps because we live and work in a time-constrained and often reward-driven culture.

- Semi-supervised learning: This method combines aspects of supervised and unsupervised learning, where the process and reference data are known, along with the intended result—but the data is incomplete. It pulls from supervised learning using the reference data that's available, but also incorporates unsupervised learning to make a best guess on the new result.
- **Reinforcement learning:** This method uses rewards and feedback to find the optimal method of accomplishing a task. For example, this is used to train robots with artificial intelligence (AI), and we experience it when playing video games that give rewards after completion of a task.

Conclusion:

So, it may be observed that Machine Learning and Deep Learning techniques are being extensively researched for their employment within the field of Natural Language Processing. These learning techniques are playing a crucial role in the majority of the processing of natural language tasks yet as in most applications of natural language processing. All the various processing of natural language tasks and therefore the different applications of natural language processing are different fields of research by themselves. Machine learning might be a good solution for analyzing text data. It's vital purely rules- based text analytics is dead-end. But it is not enough to use a single type of machine learning model. Certain aspects of machine learning are very subjective. We need to tune or train your system to match your perspective. In conclusion, it can be said that Machine Learning and Deep Learning techniques have been playing a very positive role in Natural Language Processing and its applications.

REFERENCES

- 1. Tatwadarshi P. Nagarhalli, Dr. Vinod Vaze, and Dr. N. K. Rana, "Impact of Machine Learning in Natural Language Processing: A Review", Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021), 2021.
- 2. Aravind Pai, What is Tokenization in NLP? Here's All You Need To Know. Available at: https://www.analyticsvidhya.com/blog/2020/05/ what-is-tokenization-nlp/.
- 3. F. Jia, Y. Lei, J. Lin, X. Zhou, N. Lu, Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, Mech. Syst. Sig. Process. 72 (2016) 303–315.4.

22

Vol. 14, Issue. 2, No. 2: 2023

- 4. 4. D.H. Milanez, L.I.L. de Faria, R.M. do Amaral, J.A.R. Gregolin, Claim-based patent indicators: a novel approach to analyze patent content andmonitor technological advances, World Patent Info. 50 (2017) 64–72.
- Poli, M. Healy, A. Kameas, Theory and Applications of Ontology: Computer Applications, Springer, Berlin, 2010.Cite this article as :Anuj Kumar Dwivedi, Mani Dwivedi, "A Study on The Role of Machine Learning in Natural Language Processing", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 4, pp. 192-198, July-August 2022.