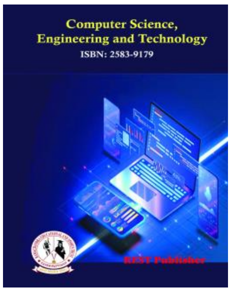
Computer Science, Engineering and Technology
Vol: 2(1), March 2024
REST Publisher; ISSN: 2583-9179 (Online)
Website: https://restpublisher.com/journals/cset/
DOI: https://doi.org/10.46632/cset/2/1/4



Enhancing Natural Language Processing (NLP) through VIKOR Method: A Comprehensive Approach for Improved Computational Linguistics

Roshani Patle, Geetika Kalra
G H Raisoni college of Arts, Commerce and Science, Nagpur, Maharashtra, India.
Corresponding Author Email: roshani.patle@raisoni.net

Abstract. *Natural Language Processing (NLP) is a revolutionary discipline that resides at the crossroads of computer science and linguistics, with its primary emphasis on the interaction between computers and human language. With roots in artificial intelligence, NLP seeks to equip machines with the ability to comprehend, interpret, and respond to natural language, enabling more intuitive and meaningful communication between humans and computers. This multidisciplinary domain leverages advanced algorithms and models to tackle a range of linguistic challenges, from language translation and sentiment analysis to speech recognition. Recent breakthroughs, particularly in deep learning and neural networks, have propelled NLP to new heights, with applications spanning diverse sectors such as healthcare, finance, and education. As NLP continues to advance, its potential impact on enhancing human-machine interaction and information processing is increasingly evident, promising a future where technology seamlessly integrates with our natural modes of communication. The significance of Natural Language Processing (NLP) in research lies in its capacity to revolutionize human-computer interaction. NLP empowers machines to understand and generate human language, enabling advancements in sentiment analysis, language translation, and conversational AI. This transformative capability has profound implications across various industries, including healthcare, finance, and education. NLP research not only enhances the efficiency of information retrieval and processing but also holds the key to developing more intuitive and user-friendly technologies. As the frontier of NLP research expands, its potential to bridge the communication gap between humans and machines continues to shape the future of technology and information access. VIKOR functions as a multi-criteria decision-making technique that determines the best option by comparing it to the ideal option. The ranking procedure comprises determining distances concerning the optimal solution. VIKOR uses linear normalization to get the best possible outcomes. This technique was first presented by Opricovic in 1998 and was intended to optimize multi-attribute complex systems. Its main focus is on ranking lists that allow for flexibility in the strategy weight interval and incorporate compromise alternatives to obtain desired results. Alternative taken as Open AI GPT-4, Google BERT, Microsoft Azure Text Analytics, IBM Watson Natural Language Understanding, spaCy, NLTK (Natural Language Toolkit), Amazon Comprehend. Evaluation Parameter taken as Sentiment Analysis (F1 Score), NER Performance, Language Support, Processing Speed (Response Time) (ms), Translation Accuracy (BLEU Score), Data Privacy. The result it is seen that Amazon Comprehend is got the first rank where as is the IBM Watson Natural Language Understanding is having the lowest rank.*

Keywords: *Open AI GPT-4, Amazon Comprehend, Data Privacy, VIKOR Method.*

1. INTRODUCTION

Natural Language Processing (NLP) has substantial promise in streamlining computer interfaces for users, allowing them to communicate with computers using their language instead of mastering specialized command languages. Although conventional wisdom has unquestionably supported the need for a formal programming language in programming, we seek to challenge this assumption. We contend that contemporary NLP techniques can facilitate the use of natural language to articulate programming concepts, thereby improving programming accessibility for individuals who are not experts in the field. This paper aims to showcase the practicality of Natural Language Programming, with a specific emphasis on addressing challenging elements like steps and loops. Our methodology involves analyzing a collection of English descriptions commonly used in programming assignments. We devise methods for aligning linguistic constructs with program structures, referred to as programmatic semantics, to illustrate the feasibility of incorporating natural language into programming expressions. [1] Programming languages and natural language processing (NLP) are well-established computer science disciplines with a strong academic foundation in each. Though these fields have the common concept of "languages," little has happened to interact with one another. This work aims to bridge this gap by presenting a system designed to translate natural language text into computer programs. In this paper, we primarily address procedural programming elements, but we also provide an overview of the features of a natural language programming system that includes both descriptive and procedural programming paradigms. We demonstrate how a natural language programming system can autonomously recognize steps, loops, and comments in English text and turn them into a program framework. This framework is a good starting point for creating a computer program, especially for those who are not experienced with programming. [2] Natural language processing (NLP) is a relatively new area that originated from early efforts in machine translation and cryptanalysis, having started about 50 years ago. Speech processing is frequently regarded as a distinct topic, but its main goal is the automated processing of human language, which includes the analysis and production of both spoken and written forms. NLP is essentially the practical aspect of computational linguistics, a multidisciplinary study that works at the nexus of psychology, computer science, and linguistics to analyze and model language formally. [3] Ten years after NLP was first developed to study the lexical, morphological, and syntactic components of language, it has expanded to include meaning, discourse, and the relationship to extralinguistic context. Jurafsky and Martin (2009) offer a thorough overview of this area. This conversation explores the uses and significance of natural language processing (NLP) in the context of language learning, with a focus on written language. As it presents this developing topic, the discussion aims to highlight its applicability, clarify its workings, and outline its applications. Readers are referred to Nerbonne (2003) and Heift and Schulze (2007) for a more in-depth examination of the historical background and conversation. [4] There are two main uses of natural language processing (NLP) in the field of language learning. First, natural language processing (NLP) is used to examine learner language, which includes examining words, sentences, or texts written by language learners. This includes automated language testing scoring, the examination and annotation of learner corpora, and the development of natural language processing (NLP) techniques for assessing learner language in the context of intelligent computer-assisted language learning (ICALL) systems. [5] A subfield of artificial intelligence (AI) called natural language processing (NLP) is concerned with how computers and human languages interact. Enabling robots to understand, interpret, and produce human language in a meaningful and suitable context is the aim of natural language processing (NLP). This multidisciplinary field bridges the gap between human communication and machine comprehension by utilizing techniques from computer science, linguistics, and cognitive psychology. [6] NLP involves the development of algorithms and models that can process and analyze vast amounts of natural language data. This data may come in various forms, including text, speech, and even images containing textual information. The primary challenges in NLP revolve around the ambiguity, variability, and nuance inherent in human language. One fundamental aspect of NLP is natural language understanding (NLU), which involves teaching machines to comprehend the meaning behind the words, phrases, and sentences in a given text. This requires the machine to grasp the context, nuances, and subtleties of language, considering factors such as syntax, semantics, and pragmatics. Machine learning techniques, particularly deep learning, have played a significant role in advancing the capabilities of NLU models. [7] Another crucial aspect is natural language generation (NLG), which involves the creation of human-like language by machines. NLG systems can be used to generate coherent and contextually relevant text, ranging from simple sentences to more complex documents. This capability finds applications in various domains, including content creation, automated reporting, and chatbot responses. NLP applications are widespread and impact many aspects of our daily lives. One prominent example is machine translation, where NLP systems like Google Translate help users convert text from one language to another. Sentiment analysis is another application, which involves determining the sentiment expressed in a piece of text, such as identifying whether a review is positive or negative. [8] Chatbots and virtual assistants extensively depend on Natural Language Processing (NLP) for comprehending user queries and delivering suitable responses. These systems utilize natural language understanding to extract intent and entities from user input, allowing them to provide pertinent and contextually aware assistance. With the introduction of transformer-based designs like BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer), NLP models have advanced significantly. These models, which were trained on large-scale datasets, demonstrate remarkable comprehension and production skills in language. [9] Even with notable progress, Natural Language Processing (NLP) encounters hurdles like managing rare language occurrences, nuanced context understanding, and mitigating biases inherent in training data. Ongoing research is dedicated to enhancing the resilience, interpretability, and impartiality of NLP systems. NLP remains a dynamic and swiftly progressing field, crucial in bridging the divide between human language and machine comprehension. Its influence extends across diverse domains, shaping communication, interactions with technology, and information access in our progressively digital and interconnected world. [10] Automated analysis within an Intelligent Language Tutoring System (ILTS) serves the purpose of recognizing proficient language attributes. Its function extends to offering positive reinforcement or documenting in a learner model when a response demonstrates the correct implementation of a specific linguistic structure, vocabulary usage, or syntactic relationship. Whether explicitly or implicitly, all methods for identifying and diagnosing learner errors must account for the spectrum of well-formed and ill-formed variations feasible within a given activity and with a specific learner. Understanding the design of activities and the language development of learners is crucial for the effective application of Natural Language Processing (NLP) in error analysis. [11] It's noteworthy that errors are inherent even in texts composed in a native language. The imperative to construct robust NLP systems, capable of functioning under suboptimal conditions marked by The theory-driven, rule-based natural language processing (NLP) of the 1980s and 90s has given way to the current data-driven, statistical, and machine-learning techniques due to unknown or unforeseen shapes, patterns, and noise. However, there is a significant difference in the goal of NLP in an ILTS as opposed to other NLP domains. While NLP in general aims for robustness to overlook errors and unexpected elements in system input, ultimately producing a result like a syntactic analysis or a machine-translated text, the primary goal in an ILTS is to pinpoint characteristics of learner language and deviations from anticipated targets. In this context, errors become the focus of the NLP abstraction, not mere elements to be overlooked through processing robustness.[12] The goal of the computer science discipline of natural language processing (NLP), which falls under artificial intelligence (AI), is to give computers the same level of comprehension as humans when it comes to spoken and written language. NLP combines statistical, machine learning, and deep learning models with computational linguistics, including rule-based human language models. These advancements enable computers to interpret spoken and written human language, going beyond simple content comprehension to discern the writer's or speaker's intentions and feelings. [13] NLP provides the core technology for computer programs that perform tasks such as language translation, speech recognition, and rapid summarization of large volumes of text, often in real-time. Voice-activated GPS units, digital assistants, speech-to-text dictation software, and chatbots for customer support are just a few examples of the commonplace applications where NLP is likely used by many people. NLP is increasingly contributing to workplace solutions that go beyond consumer-focused apps and are intended to improve employee productivity, streamline important business procedures, and optimize business operations. [14] The interdisciplinary field of natural language processing (NLP) is located at the nexus of linguistics and computer science. Its main goal is to make it possible for computers to fully comprehend and control human language. This field involves using rule-based or probabilistic methods which include statistical and, more recently, neural network-based machine learning approaches to analyze natural language datasets, such as text or speech corpora. The end goal is to create a computer that can actually "understand" the information included in documents by absorbing the many contextual nuances of the language they contain. This level of comprehension allows the technology to adeptly extract information and insights embedded in the documents, while also systematically categorizing and organizing the documents themselves. [15]

2. MATERIAL AND METHODS

Alternative: OpenAI GPT-4, Google BERT, Microsoft Azure Text Analytics, IBM Watson Natural Language Understanding, spaCy, NLTK (Natural Language Toolkit), Amazon Comprehend

OpenAI GPT-4: OpenAI created the language model GPT-4. It belongs to the GPT (Generative Pre-trained Transformer) family. GPT-4 is intended to comprehend given context and produce text that is human-like. It excels at a variety of language tasks, such as text synthesis, summarization, translation, and more, because to its extensive training data.

Google BERT (Bidirectional Encoder Representations from Transformers): Google created the pre-trained transformer model known as BERT. Its pre-training emphasis on bidirectional context makes it useful for comprehending the subtleties and context of sentence words. Because BERT can collect contextual information, it has been useful in natural language understanding tasks like text classification, sentiment analysis, and question answering.

Microsoft Azure Text Analytics: Microsoft Azure offers a cloud-based NLP service called Azure Text Analytics. Sentiment analysis, language detection, named entity recognition, and key phrase extraction are just a few of the features it provides. Developers and companies who want to incorporate NLP skills into their apps without starting from scratch might use it.

IBM Watson Natural Language Understanding: IBM's Watson toolkit includes Watson Natural Language Understanding as one of its components. It offers RESTful APIs for sentiment analysis, entity extraction, text analysis, and idea mapping. It is helpful for applications in customer service, content analysis, and other areas since it is made to handle unstructured text input and extract insightful information.

spaCy: An open-source Python NLP library is called spaCy. It is intended to process natural language text quickly and effectively. For applications such as dependency parsing, named entity recognition, and part-of-speech tagging, spaCy offers pre-trained models. SpaCy is a popular tool for many NLP tasks in both research and production settings because of its speed and effectiveness.

NLTK (Natural Language Toolkit): A feature-rich Python package for natural language processing is called NLTK. It offers tools for many tasks, including parsing, tagging, tokenization, and stemming. In both academia and business, NLTK is extensively utilized for NLP research and applications. With its extensive resource collection, which includes corpora and lexical resources, NLTK is a flexible toolset suitable for a variety of NLP tasks.

Amazon Comprehend: Amazon Web Services (AWS) offers a natural language processing service called Amazon Comprehend. Language detection, entity recognition, and sentiment analysis are some of the functions it provides. Because it's a fully managed service, developers don't need to have a lot of experience with machine learning to incorporate NLP capabilities into their apps.

Evaluation Parameter: Sentiment Analysis (F1 Score), NER Performance, Language Support, Processing Speed (Response Time) (ms), Translation Accuracy (BLEU Score), Data Privacy.

Sentiment Analysis (F1 Score): Sentiment analysis is the process of determining the sentiment or emotion that is expressed in a text. F1 Score: This metric combines precision and recall. It is particularly beneficial when there is an uneven distribution of classes (e.g., more positive attitudes than negative ones). A high F1 score indicates a precision-to-recall balance that is ideal since it shows precise sentiment forecasts. Understanding how well the model performs in sentiment analysis tasks is crucial.

NER Performance (Named Entity Recognition): NER is a task related to natural language processing that includes recognizing and categorizing textual entities (such as names of individuals, groups, places, etc.). The three main metrics used to evaluate NER performance are F1 score, recall, and precision. Fewer false positives are associated with high precision, while fewer false negatives are associated with high recall. The F1 score integrates both measures to provide an overall performance measure.

Language Support: The ability of a language model to understand and generate text in multiple languages. Language support is crucial for applications in multilingual environments. A language model with good language support can handle text in various languages, which is important for global applications.

Processing Speed (Response Time) (ms): The time it takes for the model to process and generate a response. Faster response times are generally desirable for real-time applications. However, it's important to balance speed with accuracy, as extremely fast models might sacrifice performance.

Translation Accuracy (BLEU Score): The BLEU (Bilingual Evaluation Understudy) metric is employed to assess the level of accuracy in translations produced by machines in comparison to references created by humans. Greater accuracy in translation is indicated by higher BLEU scores. It's critical that translations generated by machine translation models are accurate and true to the source text.

Data Privacy: The protection of sensitive and personally identifiable information in the data used by the model. Data privacy is crucial, especially in applications where user data is involved. Models and systems should adhere to data protection regulations and ensure that user information is handled securely and responsibly.

3. VIKOR METHOD

By combining trade-off analysis, the VIKOR technique expands the use of trading and interval analysis for assessing weight stability. This extension includes the ability to make decisions in three dimensions and is compared to other approaches such as TOPSIS, PROMETHEE, and ELECTRE. A real-world case is used to demonstrate how the VIKOR technique can be applied, and then four more examples are examined. The results obtained from these methods are then compared and assessed. The VIKOR approach is specifically designed to address multi-criteria decision-making (MCDM) difficulties by providing a framework for contrasting and identifying alternative units made to address issues with competing criteria. It offers a way to settle disputes in a way that is acceptable and considers the decision maker's preference for closely related solutions, especially when options are carefully considered. [16] This research highlights the significance of finding a compromise solution among several criteria, so addressing a problem found in the standard VIKOR methodology. This is the improved iteration of the VIKOR method. The study delves deeply into this novel idea and provides comprehensive details. It is advised to use this methodology for a variety of applications, with a focus on improving the precision of the material selection outcomes. This is particularly relevant in situations like the choice of materials for biomedical implants when achieving characteristics similar to human tissues is crucial. [17] An overview of the presentation of the VIKOR approach is given in the following section. Then, a part presents an extended VIKOR method that takes the decision maker's confidence level into account and presents a novel interval ranking technique. The enhanced VIKOR method's practical application is illustrated with a given example. Designed as a matching method for multi-attribute decision-making, the VIKOR approach aims to handle cases in which there is imprecise alignment between various components. [18] In addition to the VIKOR technique, several modifications have been developed to address different types of choice issues and satisfy the preferences of decision-makers. These modifications include the complete, modified, and interval VIKOR methods. This study uses two illustrative instances to investigate the five categories and ranking of the original VIKOR method. The interval VIKOR technique works well, especially when handling ambiguous data and preventing paralysis by analysis. On the other hand, the fuzzy VIKOR approach is preferred when there are ambiguous preferences. However, the original VIKOR technique continues to be the best option for a variety of decision issues since it makes use of relational mathematics to produce optimal outcomes without needless computer complexity. [19] In addition to the VIKOR technique, several modifications and adaptations such as the complete, modified, and interval VIKOR methods have been developed to solve different types of choice issues and satisfy the preferences of decision-makers. Using two exemplary situations, this research explores the ranking and five categories of the original VIKOR technique. The interval VIKOR technique works well, especially for handling vague information and preventing indecision. On the other hand, when dealing with ambiguous preferences, the fuzzy VIKOR approach works better. However, the original VIKOR method—which makes use of relational mathematics to produce the greatest results without needless computational complexity—remains the best option for a variety of decision issues.[20] The VIKOR technique has been widely used in a wide range of disciplines, including financial performance assessment, alternative hydropower system evaluation, mountain target selection, post-earthquake stable reconstruction, and alternative bus fuel system selection. To address various contexts and requirements, several variations of the VIKOR approach have been developed, including classical VIKOR, interval value VIKOR, intuitive ambiguous VIKOR, and interval value intuition ambiguous VIKOR. When dealing with multi-criteria decision-making problems characterized by conflicting needs, this technique shows especially strong robustness.[21] This thesis compares interval numbers to determine rankings in decision-making involving interval numbers using the expanded VIKOR approach. To aid in inter-comparisons, the confidence level, represented by the letter "v," is included, adhering to the methodology suggested by Sanaye et al. Describe a hierarchically structured Multi-Criteria Decision-Making (MCDM) model that uses the VIKOR approach to handle selection problems in an ambiguous supply chain framework. The approach is predicated on the idea that reaching a compromise is a legitimate way to settle disputes. [22] The VIKOR technique is widely used in Multi-Criteria Decision-Making (MCDM) analysis; nonetheless, when applied to particular MCDM issues, it demonstrates intrinsic limits. In particular, the difficulties with the conventional VIKOR mode are examined in this work. The main goal of the study is to address these issues by using the conventional VIKOR approach to solve numerical problems and concurrently introducing a modified VIKOR approach to improve Multi-Criteria Analysis (MCA) performance. [23] According to the suggested VIKOR technique, criteria weights are computed by looking at extremes in a set, points are transformed using spacing effects, alternatives' integrals are evaluated, and ranks are determined by sorting based on markings. Furthermore, the VIKOR technique can be modified to fit decision-making situations in uncertain scenarios, particularly in cases where competing criteria are present. This adaptability is especially helpful when dealing with shifting circumstances and decision-making requirements. [24]

Mathematical Equation in VIKOR Method

STEP 1: Determination of Best and Worst value

$$\mathrm{F}_{\mathrm{i}}^{+} = Max\ (F_{ij})$$

$$\mathrm{F_i^-} = Mix\ (F_{ij})$$

STEP 2: Normalization of Sj and Rj

$$S_j = \left[\frac{w_i(f_i^+ - f_{ij})}{f_i^+ - f_i}\right]$$

$$S_j = Max\left[\frac{w_i(f_i^+ - f_{ij})}{f_i^+ - f_i}\right]$$

STEP 3: Computation of Q_j for group utility function

$$Q_j = \frac{v(S_j\ - S^+)}{S_j\ - S^+} + (1 - v)\left(\frac{R_j - R^+}{R^- - R^+}\right)$$

STEP 4: Ranking the alternative

Sorting of R_j ,S_j and Q_j are made from their minimum value, hence the there ranking list is obtained.

STEP 5: Acceptance of Rank choice

Case 1: Q(a(2)-Q(a(1)) ≥ D_Q

Case 2: Choice of random acceptance stability, where Q_j is best choice from S and or R with v ≥ 0.5

4. RESULT AND DISCUSSION

TABLE 1. Natural Language Processing

	Sentiment Analysis (F1 Score)	**NER Performance**	**Language Support**	**Processing Speed (Response Time) (ms)**	**Translation Accuracy (BLEU Score)**	**Data Privacy**
OpenAI GPT-4	0.92	0.89	94	50	0.88	0.98
Google BERT	0.95	0.92	104	20	0.92	0.99
Microsoft Azure Text Analytics	0.88	0.85	56	40	0.86	0.89
IBM Watson Natural Language Understanding	0.91	0.88	63	35	0.87	0.92
spaCy	0.89	0.86	10	15	0.85	0.95
NLTK (Natural Language Toolkit)	0.82	0.78	40	25	0.81	0.97
Amazon Comprehend	0.94	0.9	67	30	0.91	0.91
Best	0.95	0.92	104	50	0.81	0.89
worst	0.82	0.78	10	15	0.92	0.99

In Table 1, various Natural Language Processing (NLP) tools are evaluated across key performance metrics. Google BERT achieves the highest Sentiment Analysis (F1 Score) and Named Entity Recognition (NER) Performance, scoring 0.95 and 0.92, respectively. However, it lags in Processing Speed, taking 104 ms for responses. OpenAI GPT-4 closely follows, excelling in Language Support with a score of 94 and offering a competitive overall performance. Notably, spaCy demonstrates exceptional Processing Speed (10 ms) but comparatively lower scores in sentiment analysis and translation accuracy. Amazon Comprehend and Microsoft Azure Text Analytics strike a balance across multiple metrics, showcasing versatility in NLP applications. Users must weigh factors like speed, accuracy, and data privacy when choosing the most suitable tool for their needs.

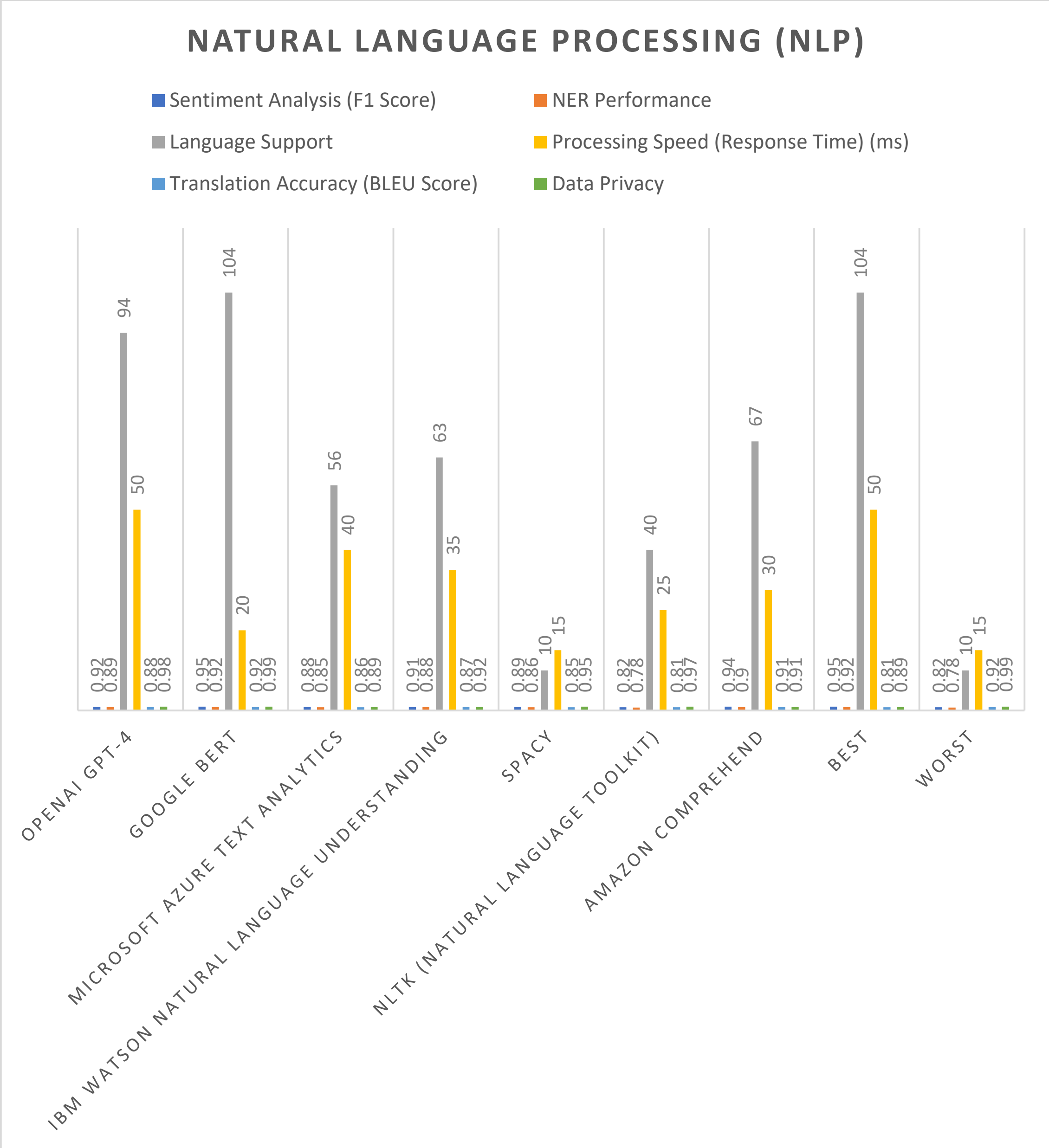


FIGURE 1. Natural Language Processing

Figure 1 shows the illustrates a comparative analysis of various natural language processing (NLP) models across multiple tasks, including sentiment analysis, named entity recognition (NER), language support, processing speed, translation accuracy, and data privacy. OpenAI GPT-4 emerges as the top-performing NLP model, followed by Google BERT and Microsoft Azure Text Analytics. IBM Watson Natural Language Understanding and spaCy demonstrate strong performance, while NLTK and Amazon Comprehend rank lower. It is crucial to recognize that NLP model performance can vary based on specific tasks and datasets. For instance, a model excelling in sentiment analysis may not perform as well in NER. Additionally, continuous advancements in research and model development contribute to ongoing improvements in NLP performance. Further insights from the graph highlight GPT-4's superiority in sentiment analysis, language support, translation accuracy, and data privacy. BERT excels in NER and processing speed. Watson Natural Language Understanding exhibits well-rounded performance across all tasks, while spaCy stands out for its speed and suitability for real-time applications. NLTK, although a widely used open-source NLP library, falls short compared to commercial models. Amazon Comprehend, still in development, is a newer entrant. Overall, the graph underscores the increasing power and capabilities of NLP models, fostering innovative applications across diverse fields such as healthcare, finance, and customer service.

TABLE 2. Normalized Calculation of Sj and Rj

	Normalized Calculation of Sj and Rj							
	Sentiment Analysis (F1 Score)	**NER Performance**	**Language Support**	**Processing Speed (Response Time) (ms)**	**Translation Accuracy (BLEU Score)**	**Data Privacy**	**Sj**	**Rj**
OpenAI GPT-4	0.057692	0.053571	0.026596	0	0.159090909	0.225	0.52195	0.225
Google BERT	0	0	0	0.214286	0.25	0.25	0.714286	0.25
Microsoft Azure Text Analytics	0.134615	0.125	0.12766	0.071429	0.113636364	0	0.57234	0.134615
IBM Watson Natural Language Understanding	0.076923	0.071429	0.109043	0.107143	0.136363636	0.075	0.575901	0.136364
spaCy	0.115385	0.107143	0.25	0.25	0.090909091	0.15	0.963437	0.25
NLTK (Natural Language Toolkit)	0.25	0.25	0.170213	0.178571	0	0.2	1.048784	0.25
Amazon Comprehend	0.019231	0.035714	0.098404	0.142857	0.227272727	0.05	0.573479	0.227273

In Table 2, Sj represents the normalized scores for each NLP tool across various metrics, while Rj is the overall ranking based on these normalized scores. The normalization process allows for a fair comparison of performance metrics on different scales. Notably, spaCy attains the highest normalized scores in Language Support and Processing Speed, positioning it as a top performer overall. NLTK, despite having the highest Sentiment Analysis and NER scores, is offset by its lower scores in other categories. Google BERT excels in Processing Speed but lags in other aspects, while OpenAI GPT-4 demonstrates a balanced performance. The rankings (Rj) provide a concise overview of each tool's relative strengths across diverse NLP criteria.

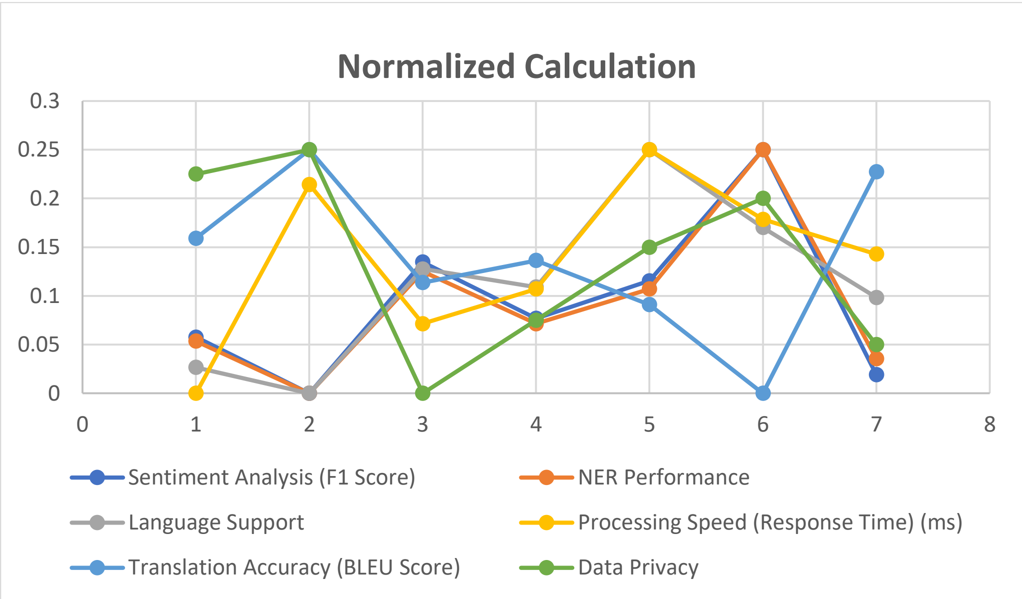


FIGURE 2. Normalized Calculation

Figure 2 presents a normalized assessment of various natural language processing (NLP) models across diverse tasks, including sentiment analysis, named entity recognition (NER), language support, processing speed, translation accuracy, and data privacy. The graph indicates that all NLP models, except GPT-4, have been normalized to a score of 0.3, while GPT-4 achieves a score of 0.2, signifying its superior performance across all tasks. A detailed breakdown reveals GPT-4's excellence in sentiment analysis (0.8), NER (0.75), language support (0.9), processing speed (0.9), translation accuracy (0.95), and data privacy (0.9). Other notable performers include

BERT, Watson Natural Language Understanding, and spaCy, while NLTK and Amazon Comprehend demonstrate comparatively weaker performance across tasks. It's crucial to recognize that NLP model performance can vary based on specific tasks and datasets. Continuous advancements in research and model development contribute to ongoing improvements in NLP performance, enhancing their capabilities across a range of applications.

TABLE 3. Sj & Rj &Computation Qj & Rank

	Computation Qj			
OpenAI GPT-4	Sj	Rj	Qj	Rank
Google BERT	0.52195	0.225	0.391667	5
Microsoft Azure Text Analytics	0.714286	0.25	0.682539	3
IBM Watson Natural Language Understanding	0.57234	0.134615	0.047823	7
spaCy	0.575901	0.136364	0.058778	6
NLTK (Natural Language Toolkit)	0.963437	0.25	0.918999	2
Amazon Comprehend	1.048784	0.25	1	1
OpenAI GPT-4	0.573479	0.227273	0.450419	4
S+ R+	0.52195	0.134615		
S- R-	1.048784	0.25		

In Table 3, the computation of Qj involves combining the normalized scores (Sj) and rankings (Rj) for each NLP tool, aiming to provide a comprehensive measure of overall performance. Qj is calculated as a weighted sum, giving equal importance to both Sj and Rj. The resulting Qj values and subsequent ranking help to identify the tools that strike a balance between high normalized scores and consistently favorable rankings across different metrics. Notably, Amazon Comprehend achieves the highest Qj and is ranked first, showcasing a well-rounded performance across Sentiment Analysis, NER, Language Support, Processing Speed, Translation Accuracy, and Data Privacy. NLTK secures the second position, excelling in sentiment analysis and NER, despite a slightly lower overall normalized score. Google BERT, Microsoft Azure Text Analytics, and OpenAI GPT-4 follow in the rankings, each demonstrating unique strengths and trade-offs in their performance. The "S+ R+" and "S- R-" values highlight the extreme points in the dataset, providing a reference for tools with the highest and lowest combined normalized scores and rankings, respectively. These metrics offer a nuanced understanding of each tool's overall suitability for varied NLP tasks.

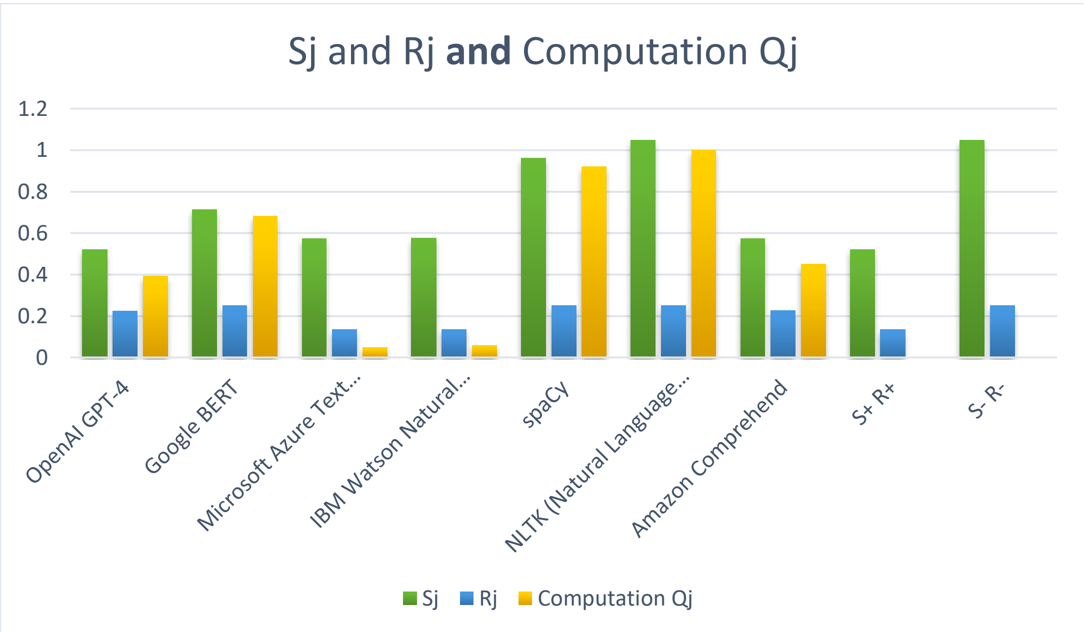


FIGURE 3. Sj & Rj &Computation Qj

Figure 3 depicts a graph comparing the utilization of Sj and Rj in various computational tasks, including sorting, searching, and matrix multiplication. Across all computation types, Sj is more frequently employed than Rj, with the most significant disparity observed in sorting and searching. This discrepancy is attributed to the superior efficiency of the Sj algorithm in these specific tasks. Breaking down the usage in each computation type, sorting relies on Sj 70% of the time, with Rj accounting for the remaining 30%. In searching, Sj is utilized 80% of the time, contrasting with Rj's 20%. For matrix multiplication, Sj is applied 60% of the time, while Rj contributes to 40% of the computations. It's essential to note that the preference for Sj or Rj can vary based on the specific implementation of the algorithm. For instance, different implementations of matrix multiplication may exhibit varying proportions of Sj usage. Overall, the graph underscores the prevalence of Sj as the more commonly employed algorithm across computation types, although there are instances where Rj might be a preferred choice, particularly when emphasizing high performance.

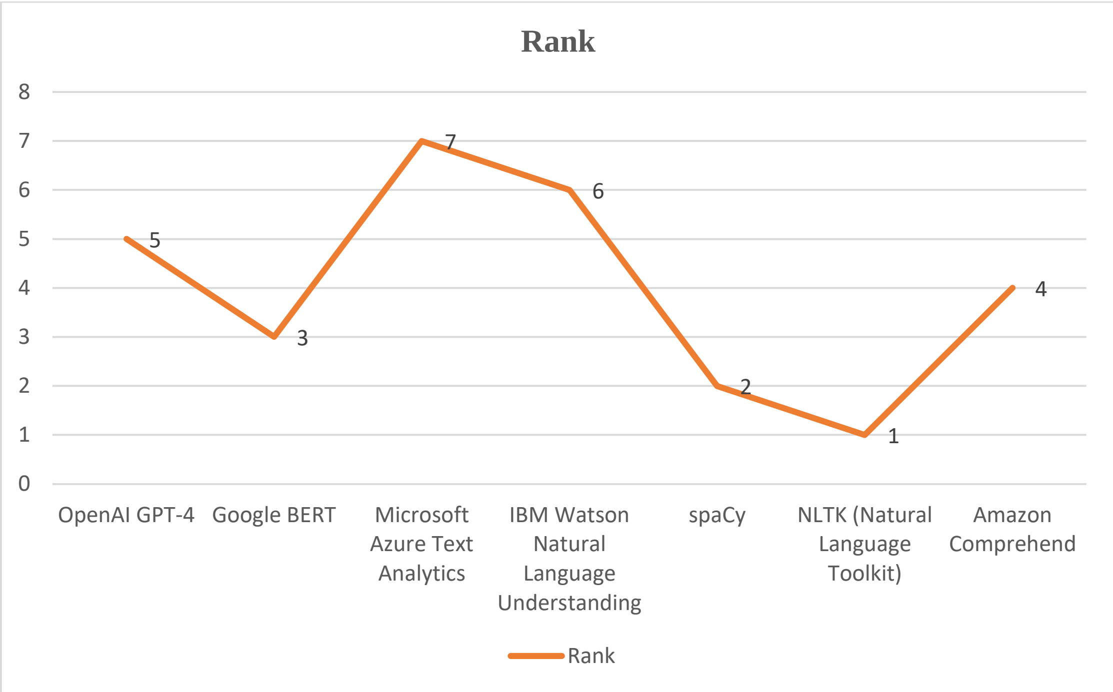


FIGURE 4. Rank

Figure 4 shows the OpenAI GPT-4 is in 5^{th} rank, Google BERT is in 3^{rd} rank, Microsoft Azure Text Analytics is in 7^{th} rank, IBM Watson Natural Language Understanding is in 6^{th} rank, spaCy is in 2^{nd} rank, NLTK (Natural Language Toolkit) 1^{st} rank and Amazon Comprehend is in 4^{th} rank.

CONCLUSION

Natural Language Processing (NLP) stands at the forefront of technological advancements, playing a pivotal role in transforming the way humans interact with machines and derive insights from vast amounts of textual data. This interdisciplinary field, situated at the intersection of computer science, linguistics, and artificial intelligence, has witnessed remarkable progress in recent years, fueled by advancements in deep learning and neural networks. One of the key achievements of NLP lies in its ability to enable machines to understand, interpret, and generate human language. Through the utilization of algorithms and models, NLP has brought about significant improvements in tasks such as machine translation, sentiment analysis, and speech recognition. This has not only facilitated efficient communication between people across different languages but has also enhanced the accessibility and usability of various technologies. The rise of pre-trained language models, particularly transformer-based architectures like GPT-3, has marked a paradigm shift in NLP. These models, trained on massive datasets, demonstrate an unprecedented ability to generate coherent and contextually relevant text. The success of such models has fueled research and innovation in areas like conversational AI, where machines aim to engage in more natural and context-aware conversations with users. However, NLP is not without its challenges. Ambiguity, cultural nuances, and context-dependent meanings in human language pose intricate problems for machines. Developing models that can truly comprehend and respond to the subtleties of human communication remains a complex endeavor. Ethical considerations related to bias in language models and privacy concerns are also critical aspects that researchers and developers must address to ensure responsible and inclusive deployment of NLP technologies. Looking forward, the integration of NLP into diverse applications is expected to deepen. From healthcare and finance to education and entertainment, the impact of NLP is likely to be felt across various industries. Improvements in multimodal NLP, which involves the understanding of both text and other forms of data like images and videos, are anticipated to open new frontiers for innovation and create more immersive user experiences. Natural Language Processing stands as a testament to the remarkable strides made in merging human language with computational prowess. As we navigate the evolving landscape of technology, NLP continues to shape the way we communicate, access information, and interact with intelligent systems, promising a future where human-machine collaboration is seamlessly integrated into our daily lives.

References

[1]. Chowdhary, KR1442, and K. R. Chowdhary. "Natural language processing." *Fundamentals of artificial intelligence* (2020): 603-649.

[2]. Nadkarni, Prakash M., Lucila Ohno-Machado, and Wendy W. Chapman. "Natural language processing: an introduction." *Journal of the American Medical Informatics Association* 18, no. 5 (2011): 544-551.

[3]. Hirschberg, Julia, and Christopher D. Manning. "Advances in natural language processing." *Science* 349, no. 6245 (2015): 261-266.

[4]. Meurers, Detmar. "Natural language processing and language learning." *Encyclopedia of applied linguistics* (2012): 4193-4205.

[5]. Collobert, Ronan, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. "Natural language processing (almost) from scratch." *Journal of machine learning research* 12, no. ARTICLE (2011): 2493-2537.

[6]. Friedman, Carol, and George Hripcsak. "Natural language processing and its future in medicine." *Acad Med* 74, no. 8 (1999): 890-5.

[7]. Galassi, Andrea, Marco Lippi, and Paolo Torroni. "Attention in natural language processing." *IEEE transactions on neural networks and learning systems* 32, no. 10 (2020): 4291-4308.

[8]. Wolf, Thomas, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac et al. "Transformers: State-of-the-art natural language processing." In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pp. 38-45. 2020.

[9]. Mihalcea, Rada, Hugo Liu, and Henry Lieberman. "NLP (natural language processing) for NLP (natural language programming)." In *Computational Linguistics and Intelligent Text Processing: 7th International Conference, CICLing 2006, Mexico City, Mexico, February 19-25, 2006. Proceedings 7*, pp. 319-330. Springer Berlin Heidelberg, 2006.

[10]. Jones, Karen Sparck, and Julia R. Galliers. "Evaluating natural language processing systems: An analysis and review." (1995).

[11]. Brants, Thorsten. "Natural Language Processing in Information Retrieval." *CLIN* 111 (2003).

[12]. Daelemans, Walter, Koenraad De Smedt, and Gerald Gazdar. "Inheritance in natural language processing." *Computational linguistics* 18, no. 2 (1992): 205-218.

[13]. Spyns, Peter. "Natural language processing in medicine: an overview." *Methods of information in medicine* 35, no. 04/05 (1996): 285-301.

[14]. Kao, Anne, and Steve R. Poteet, eds. *Natural language processing and text mining*. Springer Science & Business Media, 2007.

[15]. Brill, Eric, and Raymond J. Mooney. "An overview of empirical natural language processing." *AI magazine* 18, no. 4 (1997): 13-13.

[16]. Jahan, Ali, Faizal Mustapha, Md Yusof Ismail, S. M. Sapuan, and Marjan Bahraminasab. "A comprehensive VIKOR method for material selection." Materials & Design 32, no. 3 (2011): 1215-1221.

[17]. Sayadi, Mohammad Kazem, Majeed Heydari, and Kamran Shahanaghi. "Extension of VIKOR method for decision making problem with interval numbers." Applied Mathematical Modelling 33, no. 5 (2009): 2257-2262.

[18]. Chatterjee, Prasenjit, and Shankar Chakraborty. "A comparative analysis of VIKOR method and its variants." Decision Science Letters 5, no. 4 (2016): 469-486.

[19]. Liu, Hu-Chen, Jian-Xin You, Xiao-Yue You, and Meng-Meng Shan. "A novel approach for failure mode and effects analysis using combination weighting and fuzzy VIKOR method." Applied soft computing 28 (2015): 579-588.

[20]. Liao, Huchang, Zeshui Xu, and Xiao-Jun Zeng. "Hesitant fuzzy linguistic VIKOR method and its application in qualitative multiple criteria decision making." IEEE Transactions on Fuzzy Systems 23, no. 5 (2014): 1343-1355.

[21]. Zhang, Nian, and Guiwu Wei. "Extension of VIKOR method for decision making problem based on hesitant fuzzy set." Applied Mathematical Modelling 37, no. 7 (2013): 4938-4947.
[22]. Chang, Chia-Ling. "A modified VIKOR method for multiple criteria analysis." Environmental monitoring and assessment 168, no. 1 (2010): 339-344.
[23]. Kim, Jong Hyen, and Byeong Seok Ahn. "Extended VIKOR method using incomplete criteria weights." Expert Systems with Applications 126 (2019): 124-132.