*Asian Journal of Research in Computer Science*

*10(1): 1-17, 2021; Article no.AJRCOS.70231*
*ISSN: 2581-8260*

# Review on Natural Language Processing Based on Different Techniques

**Dastan Hussen Maulud[1*], Siddeeq Y. Ameen[1], Naaman Omar[1], Shakir Fattah Kak[1], Zryan Najat Rashid[2], Hajar Maseeh Yasin[1], Ibrahim Mahmood Ibrahim[1], Azar Abid Salih[1], Nareen O. M. Salim[1] and Dindar Mikaeel Ahmed[1]**

*[1]Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq.*
*[2]Sulaimani Polytechnic University, Sulaimani, Kurdistan Region, Iraq.*

***Authors' contributions***

*This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.*

***Article Information***

DOI: 10.9734/AJRCOS/2021/v10i130231
Editor(s):
(1) Dr. Dariusz Jacek Jakóbczak, Koszalin University of Technology, Poland.
Reviewers:
(1) Manjeet Singh, IMaharaja Ranjit Singh Punjab Technical University, India.
(2) Jay-Ar P. Lalata, FEU Institute of Technology, Philippines.
(3) Bhawna Kumari, Giani Zail Singh Punjab Technical University Campus, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/70231

***Review Article***

***Received 18 April 2021***
***Accepted 21 June 2021***
***Published 21 June 2021***

## ABSTRACT

With the exponential growth of the information on the Internet, there is a high demand for making this information readable and processable by machines. For this purpose, there is a need for the Natural Language Processing (NLP) pipeline. Natural language analysis is a tool used by computers to grasp, perceive, and control human language. This paper discusses various techniques addressed by different researchers on NLP and compares their performance. The comparison among the reviewed researches illustrated that good accuracy levels haved been achieved. Adding to that, the researches that depended on the Sentiment Analysis and ontology methods achieved small prediction error.

The syntactic analysis or parsing or syntax analysis is the third stage of the NLP as a conclusion to use NLP technology. This step aims to accurately mean or, from the text, you may state a dictionary meaning. Syntax analysis analyzes the meaning of the text in comparison with the formal grammatical rules.

**Corresponding author: E-mail: dastan.mawlud@mhe-krg.org;*

*Keywords: Natural language processing; sentiment analysis; ontology; topic modeling; tokenization; named entity recognition.*

## 1. INTRODUCTION

Natural Language Processing (NLP) [1] is the automated processing of text written in natural languages, as opposed to artificial languages such as programming languages, to attempt to decipher it [2]. It is also known as Natural Language Engineering or Inquiry Linguistics [3,4]. NLP involves various tasks, ranging from low-level tasks such as sentence segmentation to high-level tasks such as semantic annotation and opinion mining. The Semantic Web is about applying semantics, i.e., context, to data on the Web to make the web pages easier to process and manage by machines [5,6].

A core feature of the theory is introducing specific identifiers to the resources in the environment (URIs) [7,8]. Resources can be things such as "Obama" or terms such as "Politician" or relations explaining how things are linked, such as "spouse-of". NLP strategies offer a means to improve web data with semantics by inserting details about entities and relations and knowing the real-world things referenced to assign a URI to each object [9,10].

Natural language is the most potent form of communication. In computer science and artificial intelligence [11,12], voice is how machines communicate with humans (natural language). There is a wide range of algorithms, including sorting, clustering, text mining, and so on [13,14].

The urge and need to explore external space is growing more and more popular as technology advances at an increasing rate [15]. The exploration and settlement of the Moon or Mars is a major objective in the next years of international space organizations [16,17]. This means long-term exposure to microgravity, isolation, isolation and other stresses, causing harm to both their bodies and brains, under severe environment. The longer an astronaut is exposed to the space, the higher the possibility of facing psychological issues [18].

Comma Split Value (CSV) files are commonly used as a basic data format [19]. Thanks to its basic structure and creative simplicity, many of the data files that are openly released and utilized by companies are generally kept in CSV files [20,21]. However, it is extremely constrained, by the conventional keyword-making approach, to search or to recall anticipated data from CSV files that cannot describe the terms for search for or processing any search data [22].

The demand for storage space is continuously expanding with the growing quantity and diversity of logs [23]. Meanwhile the speed and precision of inquiry are becoming more and more essential in huge logs [24]. While the highly developed distributed storage technology answers mass storage and quick query problems, it is too expensive. The need for query rate is not high as logs are generated as a means to track previous action [25,26].

For more than two decades, researchers in requirement engineering (RE) have been pursuing automated manipulation of the natural language requirements, for categorization, tracing, defect identification, information extraction, and other activities [27,28]. Recent technical developments in the processing of natural languages (NLP) have led to more widespread implementation of this research in industry [29,30].

More and more individuals worldwide have begun using social media platforms such as Facebook, Twitter, Instagram etc. as Internet and social media spill over the ice [31]. As a way to communicate information worldwide, the majority of people follow Multilingualism [32]. You discuss themes over a shared platform to talk with individual speakers or a group of speakers using several languages [31]. This makes the context more complicated to grasp and makes different natural language processing (NLP) tasks much more difficult to deal with [33]. Such user conduct of multi-lingual mixing in one discussion subject is referred to as code switching with multiple community inclusion [34].

The current area of health is characterized by many, diverse, very active and geographically scattered sources of information [35,36]. In addition, the growing usage of digital health data, such as eHRs, has led to a large quantity of information being stored without precedent [37-39]. Managing this enormous quantity of data may often lead to information overload problems, with possible detrimental effects on clinical activity such as failure, delays and overall patient safety [40]. To examine data with high velocity,

volume and variability, innovative methodologies, approaches and infrastructures are required [41,42].

Question Response (QA) is one of the most common tasks of language input in natural language processing (NLP). Semantic parsing is a widely used method for understanding the natural language. Synchronous frameworks are commonly used for developing syntactic and semantic text structures [43].

The bulk of areas benefit from advanced technology and progress. In fields linked to machine learning and data mining, many applications were created [44]. But more human interaction to examine things is still required in our server subsystems [45]. Therefore, a method was presented for using data analysis and machine learning approaches to make basic servers into a smart server [46,47]. There are several tools that can monitor a number of automated systems, however the automation system can not respond how my system has been more affected by mistakes [48,49]. How can my system be loaded and load-free? Through system analysis, all these answers may be reached [50]. The term "analysis" may therefore be characterized as the process to analyze complicated structures to provide meaningful results to construct smart and robust systems [51,52]. A tool would be needed to calculate such queries when the data explosion is really complicated presently Although many tools and techniques are available for analysis and their scheme can't manage the vast data or many data and automation scripts can't offer optimum results, therefore new ways to improve optimization are essential [53].

Due to the increasing use of technology, the number of data created everyday has increased exponentially in recent years [54]. Easy access to this data in this day and age is really important [55]. While it was sufficient in the past to use structured query languages to search data stored in the hard drive, natural language is more desirable for accessing the data [56].

PART-OF-SPEECH (POS) tagging is carried out as a preprocessing step in many natural language applications. Different approaches have been tried to make POS tagging accurate. But little effort has been done in Indian languages [57,58].

In every sector of human contact, natural language processing is employed [59]. Although there are confined and complicated database requests, the written, tested, optimised and execution of structured query language declarations are still found to be difficult and devoted human resources [60]. It's hard, time consuming and too many times imprecise. Such problems can be avoided if standard processes create queries dynamically [61].

Recently, concentrate on the system needs and the manner in which they may be taken and analyzed to identify the system infrastructure on which they are dependent in the rest of the system construction [62]. Data gathered from the users to convert them to UML diagrams were tough to interpret and analyze. Some difficulty was found [63].

The article gives an example of how deep learning approaches are applied in practice to language processing and modelling [64,65]. Development of statistic models of speech aids in the prediction and improvement of the processing of speech and speech recognition of a number of recognized words and phonemes [66]. However, language modelling is currently changing to neural networks and deep learning approaches from the methodologies of statistical language modelling [67].

As internet information, instruments and procedures for automated document synthesis are more available. This is necessary [68]. Text synthesis is an important research subject in natural language processing at now. There are several ways to summarizing text [69]. We suggested to summarize the Myanmar text by Latent Semantic Analysis (LSA). Latent semántic analysis (LSA) is a natural language approach which may examine the links between documents and words by creating a number of documents and terms-related ideas [70]. The technique is unattended and does not require transmission or external knowledge. In Myanmar language there is no LSA sentence extraction. This is Myanmar's first LSA Text Summarizer [71].

As a majority of online Internet networking, Opinion Mining has become a key technique to investigate such an enormous quantity of information [72]. In a wide range of current domains, many applications appear. In the interval, viewpoints have different pronouncements that lead to research problems. The problems of study have lately turned opinion mining into a dynamic research region [73].

In natural language processing, the task of forecasting is to identify the missing characters, letters, words, expressions or phrases that can occur in a certain book fragment [74]. Several frameworks have been created since the start of the NLP using several approaches for different dialects [75]. One of the major difficulties of natural linguistic processing is the lack of content prediction [76]. In addition, most activities connected to text prediction are done in many dialects, but not in Malayalam [77]. In recent years Malayalam lacked many of the benefits that any other digital language process can offer, even though it has some of its top masterpieces, historic records and more to keep the globe interested [78].

The categorization of security requirements is crucial for the Software Engineering community to create safe, resilient and attack-compatible software [79]. This categorization allows an appropriate examination of safety needs so the development process includes suitable security systems [80]. The classification of security requirements utilized for the purpose of helping ensure that suitable safety mechanisms are created in accordance with the classification of security requirements to minimize the danger of security being introduced into late stages of development [81]. These machine learning approaches have still been proven to have issues, such as handcrafts, overfitting and poor performance with high dimensional information [82].

In the scientific realm, computers become a key component. An algorithmic method addresses the real-life challenges. Algorithms are programming free and may be constructed using any natural language with which an individual is comfortable. The difficulty, though, is how it is implemented.

This paper is organized as follows: Section I. Introduction. NLP techniques and Theoretical background are presented in section II. Assessment and recommendation are discussed in section III. IV. Describes conclusion.

## 2. NLP TECHNIQUES

The field of computer science and artificial intelligence (NLP) is a field related to computer-human interaction in a natural language. The NLP's ultimate objective is to enable computers comprehend both the language and ourselves. It is the motor behind such things as virtual assistants, language recognition, sentiment analysis, autonomous summary text, machine translation and many other things. In this article, we will discuss the principles of natural language processing, discover how NLP has benefitted in recent developments in deep learning and dig into some of its approaches.

### 2.1 Sentiment Analysis

Utilizing NLP techniques, the sentiment analysis field evaluates users' emotions and associates them with the given information. Culture may affect this area in different ways. This could be misinterpreted if it has been taken too literally. “This new gadget is bad!” Although it is evident that the title alludes to the user’s dislike of the gadget, the title might endorse the gadget to a particular age group of the community [83,84]. The sentiment analysis will determine the time at which you express your opinion [85,86]. To gather statements on a time axis can provide a better insight into peoples' feelings. Facebook and Twitter both provide challenges and opportunities for social movements in Lebanon. On the positive side, it allows people to express and express themselves freely.

The records can be carefully observed for a specified time to study trends. The data will provide a preponderance of the evidence that supports the researcher's hypothesis. With the advent of the Internet, many research fields have chosen to gather data from the Web. Companies like Google, YouTube, and Amazon know how to customize the content for the customer's best interests [87]. Depending on objective metrics such as social media likes, the number of consumers, and sales. So there is little data to study this topic. This is a challenging challenge because of a) use of different languages on one topic or blog, b) use of non-standard words that cannot be found in a dictionary, and c) emojis and symbols. These are questions relevant to both the emotion and sentiment analysis domain [1]. There is a need to provide both the social scientists and psychiatrists the necessary vocabulary and tools to analyze the Web's content and get the necessary data.

Su-Zhen Wang, et al. [88]. The presented methodology allows a thorough study of semantic corpus development technologies and proposes a new web page de-duplication algorithm based on word vector distance and TF-IDF. The thesis suggests a framework for creating a semantic corpus under a cloud-based

service. A new web page elimination algorithm is proposed based on TF-IDF and word vector distance, which can more efficiently minimize instances of replicated pages. The suggested approach filters out non-repeating webpages using TF-IDF and word vector distance algorithms on the stored corpus in Hbase. The proposed approach allows speech recognition to store information.

Shane Peeler, et al. [89]. This article advances previous work on building natural-language query interfaces to online data stores by including transitive verbs and prepositional sentences, as well as an approach for accommodating queries that involve chained complex verbs. The feasibility of this approach was illustrated when a query interface was constructed to a repository containing thousands of facts. Furthermore, the method can be employed within a relational database and describe the n-ary relations formed by transitive verbs. We are currently working on developing our application so that it can access massive data stores.

Wendi Li, et al. [90]. The developed technique explains the natural language semantic comprehension algorithm and study suitable for automated query answering through preprocessing technology, word meaning disambiguation, semantic reputation analysis, comparison tests, and performance review. It lays the foundations for further scientific research. They have preprocessed the text, analyzing the interaction and similarity of terms through algorithms so that the algorithm results can reflect logical relationships through inference. However, there are limitations to precision and versatility.

Mengzhe Li, et al. [91]. Provided a tool for systematically categorizing 300 social media messages made by transgender people in positive, negative, and neutral states. Five machine learning models used to create sentiment analysis classifiers are employed in the process. They cluster the terms in logistic regression by visualizing the word's distributions to discuss the essential words that lead to classification. The top 20 words are visualized to a word their significance is distinctly visible. In addition to this, the comment segment discusses the sense of dysphoria. The annotations indicate a high degree of consensus (Cohen's Kappa = 0.8) in all three groups. Long Short-Term Memory (LSTM) is the most effective model for classification, which has a precision of 0.85 and an overall accuracy of 0.876.

Lutfi Kerem S¸enel, et al. [92]. The approach exploits a newly proposed word embedding-based tool to produce a language similarity atlas for 76 different languages worldwide. These atlas will be used to assist researchers in selecting related language pairs or classes for linguistic purposes. Research shows that the existence of semantic similarities between two languages results from spatial similarity between those languages. Pairwise cross-lingual comparisons were quantified between 76 different languages from around the world. The findings reveal that English shares a relatively high degree of semantic similarity with most other languages. In contrast, English does not share semantic similarities with many other languages, except for Chinese.

Muhammad Taimoor Khan, et al. [93], explains how sentiment analysis works and the challenges it poses. Without addressing the NLP complexities, there will be little progress in NLP. The greatest challenge in this area is the lack of efficient and reliable resources and software. They are also proposing to improve language comprehension to improve knowledge extraction. It includes numerous social sciences and business areas. Sentiment analysis is a newer marketing tool. They used machine learning to solve the problem. They deny that NLP is worthy of answering any of its questions. Negation is based on NLU problems, including domain awareness, co-reference resolution, and word sense disambiguation. Sentiment analysis is a restricted NLP concern because it only analyzes sentiment. Complex network analysis lets you layout random text. Knowledge-based systems help increase results on particular tasks. Not all ML alternatives are proposed, but they are not the only non-ML solutions. Software data should be deducted from the textual dimension.

## 2.2 Ontology

Ontology [94] is the philosophy of being, which is concerned with ‘what is, i.e., the nature of reality and which is concerned with structural aspects of existence as such (Crotty, 1998) or whether it is possible to discover about the cosmos (Snape & Spencer, 2003) (Snape & Spencer, 2003). The SAGE Online Dictionary of Social Science Methods (2006) describes ontology as "a concept concerned with the existence of, and relationships between, different aspects of society such as social actors, cultural norms and social structures" [13]. Ontology is the study of reality and our convictions about things. Ontology

is the essence of the universe and the existence of truth, but it also determines the boundaries of what can be said about it. Bryman (2008) defines social ontology as the philosophical significance ascribed to the social world by social actors. They assume what people think is the question of deciding whether there is a truth that exists separately from human conceptions and interpretations, and whether there is a common reality or several, context-specific ones.

Kazar Okba, et al. [13]. The authors proposed using various forms of information with the aim of conversion from one language to another. The uniqueness of the suggested methods is that it seeks the contextual meaning of the word in the source language by matching the imaginary scenario the word will be found in the receiver language. Finally, they check the suggested system by checking it on actual data, and the findings are very promising, and they will be fine-tuning the system. They also used a modern approach to natural language translation. To deal with ambiguities, they have developed and implemented a translation framework based on semantic analysis, which helps us detect and handle uncertainty zones. They test the proposed method using the OWLXML file containing the grammatical and terminological information. They have built upon several proprietary tools to build a translation framework.

Mattia Atzeni, et al. [95]. Proposed a way to convert natural language instructions and requests into computer code. To resolve this challenge, they exploit the Semantic Web technology platform to create CodeOntology, an open collaboration, a collaborative framework aimed at making open source code a first-class citizen on the Internet, where it can be interlinked with other tools. Therefore, this methodology uses Code Ontology to retrieve a range of methods and code samples that are graded and combined to convert a natural language specification into a Java source code. The results show that the suggested solution is equal in efficiency to proprietary programs such as Wolfram Alpha.

Taimoor Hassan, et al. [96]. This paper presented a solution to the non-availability of semantic knowledge for improved semantic interpretation. The problem is solved by taking the representations of software specifications and annotating them with their domain-specific semantics, and assessing the impact of semantic analysis. The method using a semantic structure explicitly developed for reading and disambiguating the texts. The used architecture is based on semantic technologies that depend on the awareness of the program requirement documentation and implementation. The presented methodology reveals that adapting and combing current ontologies on the necessity of the existing environment builds the performing experiment and method that is helpful information management to software frame work.

Amin Sleimi, et al. [97], presented the semantic legal metadata that lets the reader grasp and interpret the context of legal clauses. The metadata is essential for recognition of the uniform legal specifications. However, the latest requirements engineering (RE) literature is missing guidelines on which kinds of metadata are helpful for legal requirements study. Furthermore, semantic legal metadata retrieval automation is incomplete and does not harness the full ability of Natural Language Processing. [The methods] they evaluate each of the proposed semantic metadata types and reconcile each of them. Next, we perform a qualitative analysis that will enable us to decide the types of metadata that can be collected automatically. The model proposes a harmonized conceptual model for legal requirements research and automatic extraction rules for the related metadata types. They test extraction techniques based on an individual case report. Our analysis found that the rules have correct metadata annotations. The research is positive. Whether a penalty is imposed on annotation period inaccuracies or not, they achieve accuracy between 87.4% and 97.2% and a recall between 85.5% and 94.9%.

## 2.3 Topic Modeling

The subject model is the origin of the topic model [3]. In 2009, Jo L. Hofmann developed a paradigm for natural language analysis known as Probabilistic Latent Semantic Indexing Model. In 2003, Blei and Enda produced a plan for a revised edition of the DLPA (LDA). Now it can be used as a projection model to forecast the subject distribution of papers. In this way, the subject clustering or text classification is complete according to the distribution of topics in the records that have been checked [98]. Tensor Flow is used for various things such as text analytics, machine vision, and mathematical analysis.

Xiaolong Wang, et al. [7]. Used an NLP and combined with word2vec word vector conversion technology, used a similarity calculation and the semantic analysis ability to construct an optimized LDA model, which refers to a significance sampling principle to extract topic words and use cosine similarity to evaluate repetition rate. They can effectively perform text analysis and use the Latent Dirichlet Allocation model to best optimize it. The word2vec model is used to train the model, and value sampling is used to maximize the precision and recall rate of the model. The cosine similarity is used to improve the process so the results can be done more quickly.

Monisha Kanakaraj, et al. [99]. A system is proposed that gathers data from Twitter and utilizes NLP techniques to clean the data. The ensemble techniques are applied to shape the training model. After preparation, the derived function vectors are categorized according to the model, and effects are shown. The curriculum starts with data acquisition, data analysis, and preparation, followed by classification and prediction. The core principle of the suggested solution is to improve the precision of classifications by using NLP techniques. The device in question gathers data from social networking site Twitter and does NLP techniques on such Tweets. The text is evaluated for emotion using a classification scheme.

Moreover, multiple Ensemble classification approaches are tested to decide if the data is positive, negative, or neutral. The ensemble approach outperforms the conventional classification approaches of classifying a dataset. Of the ensemble approaches, Extremely Randomized Trees achieves the highest classification rate.

Sally S.Ismail, et al. [100]. Suggested solution system for the Rich Semantic Graph to the Text module, one of three modules in ongoing research on abstractive Arabic Text Summarization. They clarified the different steps and processes that generate the module. Two problems were identified that could use the module. First, using mathematical grammatical analysis is a groundbreaking approach for Abstractive Text Summarization for the Arabic Language. Second, they should have been discussing potential phases, including selecting from different writing styles and Evaluation. They also developed the RSG reduction module and are developing the prototype for their concept, which also integrates an existing Arabic Ontology.

Jayden Sarker, et al. [101]. A new model for question answering has been developed and programmed. POS tagging and semantic sorting have been used to classify the responses from a given text. The suggested solution uses a framework for logical sorting based on lambda calculus to extract the logical types of sentences. Findings are gathered based on significant features to discover the correct answer. Unlike conventional methods, the suggested approach expands the study of syntax through lambda calculus. The proposed system for answering questions is based on a generic search methodology, which achieved an average mean accuracy of 83%, surpassing the current approach by around 11% and reaching a 95% accuracy for Yes/No questions. The proposed method is simple, graceful, and more natural than current systems. Furthermore, the proposed method obtains a higher precision in answering five forms of questions than the current system.

Lakindu Gunasekara, et al. [102]. Presents methods using semantic technologies, which NLG, and discusses the pros and cons of such structures. The research focuses on using semantic natural language generation with chatbots with lower computational costs. It can reuse it for similar domains with less coding on the natural language generation portion for small-scale level domains. The researchers provide a novel architecture for chatbot systems and a modern domain ontology for natural language learning. A natural language generation framework could be developed based on an ontology. They recommend a platform for conversational interfaces which will provide an API-based interface for text responses. The proposed solution is an API that can easily be incorporated with any device, and then a platform can be built to be used on a variety of other domains.

## 2.4 Tokenization

Tokenization[103] is the operation of separating the input text into fundamental units, called tokens, which usually correspond to words, numbers, and symbols, which are divided by white space [104]. Tokenization is a critical step in most applications involving linguistic analysis, as many sophisticated algorithms use tokens as input rather than raw text. Therefore, it is necessary to use a high-quality tokenizer, as its

failure may lead to problems in other pipeline components [105]. Each of the various kinds of tokens, such as numbers, marks, punctuation, and words, is differentiated by their capitalization degree. A sample sentence is seen in Fig. 1. Each box indicates where a token is.

Prashant Gupta, et al. [104]. The proposed method for creating an Intelligent Querying Framework (IQS) requires a user to define his questions in natural language. The framework first transforms the input sentences to a form appropriate for processing by a SQL-like script. Listing queries that are then mapped to produce the desired responses. As a result, it simplifies the research process and makes it efficient and accurate. The proposed framework provides a seamless way for non-experts to use natural language to query a database's data. QS Shows a high degree of efficiency in producing detailed and reliable queries, but it also can transform complicated queries into tractable ones. It also provides the customer with a simple method for accessing the information they need.

Artem A. Maksutov, et al. [105]. Authors utilize NLP to explain non-standard data and take these results to build an information base. An article discusses a graph database that displays relationships between different pieces of text according to data patterns. It offers ways to store text information. It helps explain the relationships between words, phrases, and sentences. The two algorithms are accurate, and performance depends on input results. Clean source text is essential for plagiarism checks. A performance graph displays errors and losses of information. Dependency parsing is the central algorithm. On specific grammatical sentences, grammar-based algorithms work better than other algorithms. We can apply these improvements to other systems. A parser can be selected based on the complexity of the input sentence. A technique for identifying relationships is used to remove key details and relationships between entities in a text. It helps create graphs between named entities.

## 2.5 Named Entity Recognition

Knowledge extraction is the method of translating information in unstructured text into easily controlled classes. Essential to this is the role of named object identification and classification (NERC) [106], a method that includes recognizing proper names in texts (NER) and how to apply them to predefined categories of interest (NEC) [107]. This sort of tool analyzes sentences, but NERC is about deriving meaning from the text. The conventional core collection of named individuals, formed for the shared NERC challenge, comprises Individual, Organization, Place, and Date and Time phrases, such as Barack Obama, Microsoft, New York, etc.

Kittiphong Sengloiluean, et al. [106]. Used a semantic approach to address queries using DBpedia and WordNet. The primary goal of this project was to find the best question answering method. The paper suggested methods for solving the problems of isolating named entities from the query and how to solve problems of comparisons to named entities. It also tested the consistency of answers to the questions. Via assessed a TREC question set, DBpedia, and the proposed solution, the proposed approach obtained an F-measure score of 93.43%, an average precision of 92.73%, and an average recall of 94.15%.

M. V. Sadhuram, et al. [107]. The use of discrete mathematics would answer the proposed work for constructing a framework in which the questions are asked. NLP is used depending on input from the data that comes in text, film, photograph, or audio. NLP refers to AI, which is used in the field of question answering (QA) method. Lexical chain recognition and keyword analysis identify a series of papers to answer a given query. The logic system is used to illustrate the right approach to the question. Scientists conducted the procedure with the SQUAD database. The findings reveal that the passage retrieval accuracy using TFIDF is as high as 69.69%; the average sum of the student's accurate estimates was 69.93%.

Fatima Zait, et al. [108]. The presented approach to ambiguity identification and resolution in natural language specifications is as early as possible using NLP and semantic web techniques. Therefore, we are more likely to evaluate vague terms and have a variety of alternative meanings. They developed a solution to define and overcome lexical and semantic ambiguities in specifications to increase the overall interpretation of the requirements. POS tagger can be implemented to define the fuzzy concepts in user specifications and how related data can be used to disambiguate ambiguous semantic concepts.

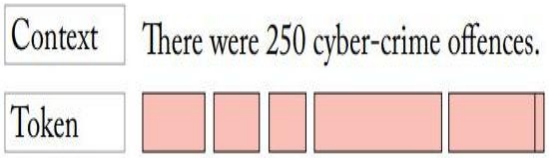


**Fig. 1. Tokenization of sentence**

## 3. ASSESSMENT AND RECOMMENDATION

Based on the previous section, researchers have adopted different approaches and techniques for various applications. However, researchers have identified the core facets of their proposals that are significant. Most of the papers included in this review are observational studies that used the NLP technique (2016 to 2020. Table 1 compares the topics explained (reviewed) in this paper. The summary includes the year of publication of each paper, technique/ method, the goal of the work, tools used in the paper, and significance and accuracy of the results. As given in Table 1, there are different techniques and methods used, Sentiment Analysis, Tokenization, ontology, Named Entity Recognition, Natural language generation, and Topic Modeling used for analyzing.

Five papers [88-92] (or 33%) are used the Sentiment Analysis technique, three papers [13,95,96] (or 20%) are used ontology technique, Named Entity Recognition, Natural language generation, and Tokenization techniques each of them is used in 2 papers [101,102,104-107] (or 13%), and one paper [7] (or 6%) is used Topic Modeling technique. Papers used different tools for implementing algorithms like Haskell, java, and OWL, but most of them used Python. Aim of the research changes depending on the techniques used, such as answering questions or analyzing and answering questions, checking for duplicate words in the language, Building interactive interfaces to annotate online documents, developing a text preprocessing system, and text translation. In all papers. The reviewed papers' best-achieved accuracy and those who depended on the Sentiment Analysis and ontology method because the Accuracy of those techniques are excellent and the prediction error is small.

From previous studies, diverse methods and strategies have been adapted. However, researchers have established the central facets of their ideas. Therefore, we suggest using the NLP technique depending on the area which new researchers are working on such as, checking word duplication and answering question systems. Sentiment Analysis is a good technique and has an excellent performance, for text translate between languages ontology technique have good efficiency and small error.

**Table 1. An overview of most recent NLP techniques**

| Ref. | Goals of the Work | Techniques | Tools | Accuracy | Other Significant Results |
|---|---|---|---|---|---|
| 2020, [89] | They are building interactive interfaces to annotate online documents. | Sentiment Analysis | Haskell, X-SAIGA | 91% | The viability of accessing NL data using an event-based semantic-web triplestore was demonstrated, which contained thousands of facts. |
| 2020, [90] | It described the natural language semantic processing of the ECA algorithm and research progress on preprocessing technology. | Sentiment Analysis | Word2Vec, GloVe | 93% | XLNet on the GLUE dataset is significantly better than BERT |
| 2020, | Check for duplicate | Topic | Python | __ | Improve the accuracy |

| Ref. | Goals of the Work | Techniques | Tools | Accuracy | Other Significant Results |
|---|---|---|---|---|---|
| [7] | words in the language. | Modeling | | | and recall rate |
| 2020, [91] | Transgender Community Sentiment Analysis From Social Media Data | Sentiment Analysis | Python | 85%. | Among the 300 Reddit comments, 72 are categorized as negative posts, while 85 are annotated to be positive. The rest, 48%, are considered as neutral sentiments. Cohen's Kappa score is over 0.8 across all classes. |
| 2020, [105] | developing text preprocessing system | Tokenization | Python | __ | All algorithms given in the solution produce correct results with a small amount of input data. |
| 2020, [107] | Answering the questions that the users ask | Named Entity Recognition | Python | 69.69%. | Overall Average over the entire dataset was 64% correct answer prediction. |
| 2019, [101] | answering questions efficiently | Natural language generation | bAbI-10k | __ | Single Supporting Facts 90%, two Supporting Facts 90%, three Supporting Facts 80%, Yes/No questions 95%, List/Sets 60%. |
| 2019, [102] | Utilize NLP techniques with semantic technologies. | Natural language generation | OWL, API | 86% | Based on impressions of 69 out of 11, the recall rate of the system is about 90%. Therefore, a system that is accurate to within 0.862% is very close to precise. |
| 2018, [92] | Construct a score of 76 different languages' similarity based on a publicly available corpus. | Sentiment Analysis | t-SNE | __ | English is in the top rank with nine other languages, while Chinese is at the bottom. |
| 2018, [95] | Translate a natural language command into an object-oriented program. | Ontology | Java | 94% | Number of Questions 120, Processed Questions 116, Correct Answers 109, Precision 0.91. |
| 2017, [13] | Translate a text from one language to another. | Ontology | OWLXM, API, java | 96% | Highly helpful! Particularly on the disambiguation and understanding measures. |
| 2017, [88] | A cloud service corpus would establish a vocabulary for representing natural language. In addition, a new webpages removal algorithm is proposed. | Sentiment Analysis | de-duplication, PageRank, Simhash, | 95% | Repeated documents 50, Detected repeated documents 41, Detected correctly repeated documents 46, Recall rate 92%. |

| Ref. | Goals of the Work | Techniques | Tools | Accuracy | Other Significant Results |
|---|---|---|---|---|---|
| 2017, [104] | A question answering system that converts questions in natural language. | Tokenization | Oracle, Microsoft SQL | 92% | Presents pragmatic ways to translate database queries into natural language that users can search. |
| 2017, [106] | It is proposed how to analyze and answer questions. | Named Entity Recognition | DBpedia, WordNet | 94.15% | The F-measure of 93.43%, an average precision of 92.73%, and over 500 questions successfully administered. |
| 2016, [96] | Formulation and analysis of natural language software requirements specifications. | Ontology | OntoGen | 71% | Software Requirements: Sample 15, Correct 10, Incorrect 04, Missing 1, Rec % 0.66 7. |

## 4. CONCLUSION

For processing some natural language, it is essential to "understand" them. The computer translation results can be vague without syntax and semantic analysis. It is a subpart of linguistics that studies meanings. This article discusses various techniques by different researchers on Natural Language Processing and compares their performance. This review utilized (15) papers in the last five years on six NLP techniques/ methods: Sentiment Analysis, Tokenization, ontology, Named Entity Recognition, Natural language generation, and Topic Modeling. The sentiment Analysis technique is suitable for checking word duplication and question answering systems. The ontology technique is suitable for Translating a text from one language to another from this study analysis. Compression results show that most papers received various degrees of accuracy, hence, helping them with their respective findings. We suggest using the NLP technique depending on the area which new researchers are working on. Sentiment Analysis is a good technique and has an outstanding performance.

Due to increased availability of raw data and cheaper computer power, the past decade saw a change in neural approaches to text interpretation. In many language problems, such approaches have proved strong and convenient. Results in distributional semantics in particular demonstrated potential approaches to capture the meaning of every word as a vector in dense low-dimensional spaces in a Text Corpus. The approach was enormously successful, as were Word2Vec, GloVe, fastText and Swivel. Recently, the area has led to the incorporation of contextual terms based on neural language models such as ELMo, BERT and GPT-2, to mention a few. Their applications include word embedding and numerous other tasks in natural language processing including semantin tagging, involvement, inquiry replies and sentimental analysis in terms of similarity, analogy and connection.

## COMPETING INTERESTS

Authors have declared that no competing interests exist.

## REFERENCES

1. Rajput A. "Natural language processing, sentiment analysis, and clinical analytics," in Innovation in Health Informatics, ed: Elsevier. 2020;79-97.
2. Zeebaree S, Ameen S, Sadeeq M. Social media networks security threats, risks and recommendation: A case study in the kurdistan region. International Journal of Innovation, Creativity and Change. 2020;13:349-365.
3. Liao X, Zhu Z. "Classification of natural language semantic relations under deep learning," in 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA). 2020;1025-1027.
4. Haji SH, Zeebaree SR, Saeed RH, Ameen SY, Shukur HM, Omar N, et al. Comparison of software defined networking with traditional networking.

Asian Journal of Research in Computer Science. 2021;1-18.

5. Rokade A, Patil B, Rajani S, Revandkar S, Shedge R. "Automated grading system using natural language processing," in 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). 2018;1123-1127.
6. Hassan RJ, Zeebaree SR, Ameen SY, Kak SF, Sadeeq MA, Ageed ZS, et al. State of art survey for iot effects on smart city technology: challenges, opportunities, and solutions. Asian Journal of Research in Computer Science. 2021;32-48.
7. Wang X, Dong X, Chen S. "Text duplicated-checking algorithm implementation based on natural language semantic analysis," in 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC). 2020;732-735.
8. Yasin HM, Zeebaree SR, Sadeeq MA, Ameen SY, Ibrahim IM, Zebari RR, et al. IoT and ICT based smart water management, monitoring and controlling system: A review. Asian Journal of Research in Computer Science. 2021;42-56.
9. Rayz JT, Raskin V. "Logic of natural language: Through the eyes of ontological semantics," in 2016 IEEE 15th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC). 2016;511-515.
10. Abdullah SMSA, Ameen SYA, Sadeeq MA, Zeebaree S. Multimodal emotion recognition using deep learning. Journal of Applied Science and Technology Trends. 2021;2:52-58.
11. Otter DW, Medina JR, Kalita JK. A survey of the usages of deep learning for natural language processing. IEEE Transactions on Neural Networks and Learning Systems; 2020.
12. Haji SH, Ameen SY. Attack and Anomaly detection in iot networks using machine learning techniques: A review. Asian Journal of Research in Computer Science. 2021;30-46.
13. Okba K, Hamza S, Hind B, Amira A, Samir B. "Semantic natural language translation based on ontologies combination," in 2017 8th International Conference on Information Technology (ICIT). 2017;315-321.
14. Izadeen GY, Ameen SY. Smart android graphical password strategy: A review. Asian Journal of Research in Computer Science. 2021;59-69.
15. Hasan BMS, Ameen SY, Hasan OMS. Image authentication based on watermarking approach. Asian Journal of Research in Computer Science. 2021;34-51.
16. Abdullah DM, Ameen SY. Enhanced mobile broadband (EMBB): A review. Journal of Information Technology and Informatics. 2021;1:13-19.
17. Khalid LF, Ameen SY. Secure Iot integration in daily lives: A review. Journal of Information Technology and Informatics. 2021;1:6-12.
18. Sharif KH, Ameen SY. "A review of security awareness approaches with special emphasis on gamification," in 2020 International Conference on Advanced Science and Engineering (ICOASE). 2020;151-156.
19. Alzakholi O, Shukur H, Zebari R, Abas S, Sadeeq M. Comparison among cloud technologies and cloud performance. Journal of Applied Science and Technology Trends. 2020;1:40-47.
20. Al Janaby AO, Al-Omary A, Ameen SY, Al-Rizzo H. Tracking and controlling high-speed vehicles via CQI in LTE-A systems. International Journal of Computing and Digital Systems. 2020;9:1109-1119.
21. Ageed Z, Mahmood MR, Sadeeq M, Abdulrazzaq MB, Dino H. Cloud computing resources impacts on heavy-load parallel processing approaches. IOSR Journal of Computer Engineering (IOSR-JCE). 2020;22:30-41.
22. Hamed ZA, Ahmed IM, Ameen SY. "Protecting windows OS against local threats without using antivirus," relation. 2020;29:64-70.
23. Mohammed K, Ameen S. Performance investigation of distributed orthogonal space-time block coding based on relay selection in wireless cooperative systems; 2020.
24. Fawzi LM, Alqarawi SM, Ameen SY, Dawood SA. Two levels alert verification technique for smart oil pipeline surveillance system (SOPSS). International Journal of Computing and Digital Systems. 2019;8:115-124.
25. Al-Sultan MR, Ameen SY, Abduallah WM. Real time implementation of stegofirewall

system. International Journal of Computing and Digital Systems. 2019;8:498-504.
26. Sallow A, Zeebaree S, Zebari R, Mahmood M, Abdulrazzaq M, Sadeeq M. Vaccine tracker. SMS reminder system: Design and implementation; 2020.
27. Al Janaby AO, Al-Omary A, Ameen SY, Al-Rizzo HM. "Tracking high-speed users using SNR-CQI mapping in LTE-A networks," in 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT). 2018;1-7.
28. Othman A, Ameen SY, Al-Rizzo H. Dynamic switching of scheduling algorithm for. International Journal of Computing and Network Technology. 2018;6.
29. Ameen SY, Ali ALSH. A comparative study for new aspects to quantum key distribution. Journal of Engineering and Sustainable Development. 2018;11:45-57.
30. Sadeeq MA, Zeebaree SR, Qashi R, Ahmed SH, Jacksi K. "Internet of things security: A survey," in 2018 International Conference on Advanced Science and Engineering (ICOASE). 2018;162-166.
31. Fawzi LM, Ameen SY, Alqaraawi SM, Dawwd SA. Embedded real-time video surveillance system based on multi-sensor and visual tracking. Appl. Math. Infor. Sci. 2018;12:345-359.
32. Ali ZA, Ameen SY. Detection and prevention cyber-attacks for smart buildings via private cloud environment. International Journal of Computing and Network Technology. 2018;6:27-33.
33. Fawzi LM, Ameen SY, Dawwd SA, Alqaraawi SM. Comparative study of Ad-hoc routing protocol for oil and gas pipelines surveillance systems. International Journal of Computing and Network Technology. 2016;4.
34. Farhan FY, Ameen SY. "Improved hybrid variable and fixed step size least mean square adaptive filter algorithm with application to time varying system identification," in 2015 10th System of Systems Engineering Conference (SoSE). 2015;94-98.
35. Othman A, Ameen SY, Al-Rizzo H. A new channel quality indicator mapping scheme for high mobility applications in LTE systems. Journal of Modeling and Simulation of Antennas and Propagation. 2015;1:38-43.
36. Al-Khayat ON, Ameen SY, Abdallah MN. WSNs Power consumption reduction using clustering and multiple access techniques. International Journal of Computer Applications. 2014;87.
37. Othman A, Othman SY, Al-Omary A, Al-Rizzo H. Comparative performanceof subcarrier schedulers in uplink LTE-A under high users' mobility. International Journal of Computing and Digital Systems. 2015;4.
38. Ameen SY, Yousif MK. Decode and forward cooperative protocol enhancement using interference cancellation. Int. J. Elect., Comput., Electron. Commun. Eng. 2014;8:273-277.
39. Ameen SY, Nourildean SW. "Coordinator and router investigation in IEEE802. 15.14 ZigBee wireless sensor network," in 2013 International Conference on Electrical Communication, Computer, Power, and Control Engineering (ICECCPCE). 2013;130-134.
40. Othman A, Ameen SY, Al-Rizzo H. An energy-efficient MIMO-based 4G LTE-A adaptive modulation and coding scheme for high mobility scenarios. International Journal of Computing and Network Technology. 2015;3.
41. Ameen SY. Advanced encryption standard (AES) enhancement using artificial neural networks. Int J of Scientific & Engineering Research. 2014;5.
42. Elzanati WM, Ameen SY. "Cost effective air-conditioning for bahrain domestic applications," in 2013 7th IEEE GCC Conference and Exhibition (GCC). 2013,;535-540.
43. Ameen SY, Nourildean SW. Firewall and VPN investigation on cloud computing performance. International Journal of Computer Science and Engineering Survey. 2014;5:15.
44. Yahia HS, Zeebaree SR, Sadeeq MA, Salim NO, Kak SF, Adel AZ, et al. Comprehensive survey for cloud computing based nature-inspired algorithms optimization scheduling. Asian Journal of Research in Computer Science. 2021;1-16.
45. Ageed ZS, Zeebaree SR, Sadeeq MM, Kak SF, Rashid ZN, Salih AA, et al. A survey of data mining implementation in smart city applications. Qubahan Academic Journal. 2021;1:91-99.

46. Ageed ZS, Zeebaree SR, Sadeeq MA, Abdulrazzaq MB, Salim BW, Salih AA, et al. A state of art survey for intelligent energy monitoring systems. Asian Journal of Research in Computer Science. 2021;46-61.
47. Ameen SY, Al-Badrany MR. "Optimal image steganography content destruction techniques," in International Conference on Systems, Control, Signal Processing and Informatics. 2013;453-457.
48. Abdulqadir HR, Zeebaree SR, Shukur HM, Sadeeq MM, Salim BW, Salih AA, et al. A study of moving from cloud computing to fog computing. Qubahan Academic Journal. 2021;1:60-70.
49. Abdulla AI, Abdulraheem AS, Salih AA, Sadeeq MA, Ahmed AJ, Ferzor BM, et al. Internet of things and smart home security. Technol. Rep. Kansai Univ. 2020;62:2465-2476.
50. Abdulraheem AS, Salih AA, Abdulla AI, Sadeeq MA, Salim NO, Abdullah H, et al. Home automation system based on IoT; 2020.
51. Salih AA, Zeebaree SR, Abdulraheem AS, Zebari RR, Sadeeq MA, Ahmed OM. Evolution of mobile wireless communication to 5G revolution. Technology Reports of Kansai University. 2020;62:2139-2151.
52. Ameen SY, Ahmed IM. "Design and implementation of e-laboratory for information security training," in 2013 Fourth International Conference on e-Learning" Best Practices in Management, Design and Development of e-Courses: Standards of Excellence and Creativity". 2013;310-317.
53. Dino HI, Zeebaree SR, Salih AA, Zebari RR, Ageed ZS, Shukur HM, et al. Impact of process execution and physical memory-spaces on OS performance; 2020.
54. Abdulrahman LM, Zeebaree SR, Kak SF, Sadeeq MA, Adel AZ, Salim BW, et al. A state of art for smart gateways issues and modification. Asian Journal of Research in Computer Science. 2021;1-13.
55. Ageed ZS, Zeebaree SR, Sadeeq MM, Kak SF, Yahia HS, Mahmood MR, et al. Comprehensive survey of big data mining approaches in cloud systems. Qubahan Academic Journal. 2021;1:29-38.
56. Yazdeen AA, Zeebaree SR, Sadeeq MM, Kak SF, Ahmed OM, Zebari RR. FPGA implementations for data encryption and decryption via concurrent and parallel computation: A review. Qubahan Academic Journal. 2021;1:8-16.
57. Malallah H, Zeebaree SR, Zebari RR, Sadeeq MA, Ageed ZS, Ibrahim IM, et al. A comprehensive study of kernel (issues and concepts) in different operating systems. Asian Journal of Research in Computer Science. 2021;16-31.
58. Ameen SY, Al-Jammas MH, Alenezi AS. "FPGA implementation of modified architecture for adaptive Viterbi decoder," in 2011 Saudi International Electronics, Communications and Photonics Conference (SIECPC). 2011;1-9.
59. Ibrahim IM. Task scheduling algorithms in cloud computing: A review. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 2021;12:1041-1053.
60. Zebari IM, Zeebaree SR, Yasin HM. "Real Time video streaming from multi-source using client-server for video distribution," in 2019 4th Scientific International Conference Najaf (SICN). 2019;109-114.
61. Yasin HM, Zeebaree SR, Zebari IM. "Arduino based automatic irrigation system: monitoring and SMS controlling," in 2019 4th Scientific International Conference Najaf (SICN). 2019;109-114.
62. Abdulazeez AM, Zeebaree SR, Sadeeq MA. Design and implementation of electronic student affairs system. Academic Journal of Nawroz University. 2018;7:66-73.
63. Zeebaree S, Yasin HM. Arduino based remote controlling for home: Power saving, security and protection. International Journal of Scientific and Engineering Research. 2014;5:266-272.
64. Zeebaree S, Zebari I. Multilevel client/server peer-to-peer video broadcasting system. International Journal of Scientific and Engineering Research. 2014;5.
65. Sallow AB, Sadeeq M, Zebari RR, Abdulrazzaq MB, Mahmood MR, Shukur HM, et al. An investigation for mobile malware behavioral and detection techniques based on android platform. IOSR Journal of Computer Engineering (IOSR-JCE). 2020;22:14-20.
66. Zebari SR, Yaseen NO. Effects of parallel processing implementation on balanced load-division depending on distributed memory systems. J. Univ. Anbar Pure Sci. 2011;5:50-56.

67. Sadeeq MM, Abdulkareem NM, Zeebaree SR, Ahmed DM, Sami AS, Zebari RR. IoT and cloud computing issues, challenges and opportunities: A review. Qubahan Academic Journal. 2021;1:1-7.
68. Kareem FQ, Zeebaree SR, Dino HI, Sadeeq MA, Rashid ZN, Hasan DA, et al. A survey of optical fiber communications: challenges and processing time influences. Asian Journal of Research in Computer Science. 2021;48-58.
69. Omer MA, Zeebaree SR, Sadeeq MA, Salim BW, Mohsin SX, Rashid ZN, et al. Efficiency of malware detection in android system: A survey. Asian Journal of Research in Computer Science. 2021;59-69.
70. Rashid ZN, Zeebaree SR, Sengur A. Novel remote parallel processing code-breaker system via cloud computing; 2020.
71. Rashid ZN, Zeebaree SR, Shengul A. "Design and analysis of proposed remote controlling distributed parallel computing system over the cloud," in 2019 International Conference on Advanced Science and Engineering (ICOASE). 2019;118-123.
72. Rashid ZN, Zebari SR, Sharif KH, Jacksi K. "Distributed cloud computing and distributed parallel computing: A review," in 2018 International Conference on Advanced Science and Engineering (ICOASE). 2018;167-172.
73. Rashid ZN, Sharif KH, Zeebaree S. Client/servers clustering effects on CPU execution-time, CPU usage and CPU idle depending on activities of parallel-processing-technique operations. Int. J. Sci. Technol. Res. 2018;7: 106-111.
74. Jijo BT, Zeebaree SR, Zebari RR, Sadeeq MA, Sallow AB, Mohsin S, et al. A comprehensive survey of 5G mm-wave technology design challenges. Asian Journal of Research in Computer Science. 2021;1-20.
75. Sadeeq MA, Zeebaree S. Energy management for internet of things via distributed systems. Journal of Applied Science and Technology Trends. 2021;2:59-71.
76. Maulud DH, Zeebaree SR, Jacksi K, Sadeeq MAM, Sharif KH. State of art for semantic analysis of natural language processing. Qubahan Academic Journal. 2021;1:21-28.
77. Shukur H, Zeebaree SR, Ahmed AJ, Zebari RR, Ahmed O, Tahir BSA, et al. A state of art survey for concurrent computation and clustering of parallel computing for distributed systems. Journal of Applied Science and Technology Trends. 2020;1:148-154.
78. Jacksi K, Ibrahim RK, Zeebaree SR, Zebari RR, Sadeeq MA. "Clustering documents based on semantic similarity using HAC and K-mean algorithms," in 2020 International Conference on Advanced Science and Engineering (ICOASE). 2020;205-210.
79. Sadeeq MA, Abdulazeez AM. "Neural networks architectures design, and applications: A review," in 2020 International Conference on Advanced Science and Engineering (ICOASE). 2020;199-204.
80. Ageed ZS, Ibrahim RK, Sadeeq MA. Unified ontology implementation of cloud computing for distributed systems. Current Journal of Applied Science and Technology. 2020;82-97.
81. Sulaiman MA, Sadeeq M, Abdulraheem AS, Abdulla AI. Analyzation study for gamification examination fields. Technol. Rep. Kansai Univ. 2020;62:2319-2328.
82. Sadeeq M, Abdulla AI, Abdulraheem AS, Ageed ZS. Impact of electronic commerce on enterprise business. Technol. Rep. Kansai Univ. 2020;62:2365-2378.
83. Hood K, Kuiper PK. "Improving student surveys with natural language processing," in 2018 Second IEEE International Conference on Robotic Computing (IRC). 2018;383-386.
84. Aziz ZAA, Ameen SYA. Air pollution monitoring using wireless sensor networks. Journal of Information Technology and Informatics. 2021;1:20-25.
85. Hadiya N, Nanavati N. "Indic senti review: Natural language processing based sentiment analysis on major Indian languages," in 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC). 2019;322-327.
86. Amanuel SVA, Ameen SYA. Device-to-device communication for 5G security: A review. Journal of Information Technology and Informatics. 2021;1:26-31.
87. Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing. Ieee

Computational intelligenCe magazine. 2018;13:55-75.
88. Wang SZ, Zhang QC, Zhang L. "Natural language semantic corpus construction based on cloud service platform," in 2017 International Conference on Machine Learning and Cybernetics (ICMLC). 2017;670-674.
89. Peelar S, Frost R. "A compositional semantics for a wide-coverage natural-language query interface to a semantic web triplestore," in 2020 IEEE 14th International Conference on Semantic Computing (ICSC). 2020;257-262.
90. Li W. "Analysis of semantic comprehension algorithms of natural language based on robot's questions and answers," in 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA). 2020;1021-1024.
91. Li M, Wang Y, Zhao Y, Li Z. Transgender community sentiment analysis from social media data: A natural language processing approach. arXiv preprint arXiv:2010.13062; 2020.
92. Şenel LK, Utlu I, Yücesoy V, Koç A, Çukur T. "Generating semantic similarity atlas for natural languages," in 2018 IEEE Spoken Language Technology Workshop (SLT). 2018;795-799.
93. Khan MT, Durrani M, Ali A, Inayat I, Khalid S, Khan KH. Sentiment analysis and the complex natural language. Complex Adaptive Systems Modeling. 2016;4:1-19.
94. Al-Saadi H. Demystifying ontology and epistemology in research methods. Research Gate. 2014;1:1-10.
95. Atzeni M, Atzori M. "Translating natural language to code: an unsupervised ontology-based approach," in 2018 IEEE first international conference on artificial intelligence and knowledge engineering (AIKE). 2018;1-8.
96. Hassan T, Hassan S, Yar MA, Younas W. "Semantic analysis of natural language software requirement," in 2016 Sixth International Conference on Innovative Computing Technology (INTECH). 2016;459-463.
97. Sleimi A, Sannier N, Sabetzadeh M, Briand L, Dann J. "Automated extraction of semantic legal metadata using natural language processing," in 2018 IEEE 26th International Requirements Engineering Conference (RE). 2018;124-135.
98. Zhang L, Tao B. A framework for ontology integration based on genetic algorithm. Journal of Intelligent and Fuzzy Systems. 2016;30:1643-1656.
99. Kanakaraj M, Guddeti RMR. "Performance analysis of Ensemble methods on Twitter sentiment analysis using NLP techniques," in Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015). 2015;169-170.
100. Ismail SS, Aref M, Moawad IF. "A model for generating Arabic text from semantic representation," in 2015 11th International Computer Engineering Conference (ICENCO). 2015;117-122.
101. Sarker J, Billah M, Al Mamun M. "Textual Question Answering for Semantic Parsing in Natural Language Processing," in 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT). 2019;1-5.
102. Gunasekara L, Vidanage K. "UniOntBot: Semantic natural language generation based API approach for chatbot communication," in 2019 National Information Technology Conference (NITC). 2019;1-8.
103. Maynard D, Bontcheva K, Augenstein I. Natural language processing for the semantic web. Synthesis Lectures on the Semantic Web: Theory and Technology. 2016;6:1-194.
104. Gupta P, Goswami A, Koul S, Sartape K. "IQS-intelligent querying system using natural language processing," in 2017 international conference of electronics, communication and aerospace technology (ICECA). 2017;410-413.
105. Maksutov AA, Zamyatovskiy VI, Vyunnikov VN, Kutuzov AV. "Knowledge base collecting using natural language processing algorithms," in 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). 2020;405-407.
106. Sengloiluean K, Arch-int N, Arch-int S, Thongkrau T. "A semantic approach for question answering using DBpedia and WordNet," in 2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE). 2017;1-6.
107. Sadhuram MV, Soni A. "Natural language processing based new approach to design factoid question answering system," in 2020 Second International Conference on

Inventive Research in Computing Applications (ICIRCA). 2020;276-281.

108. Zait F, Zarour N. "Addressing lexical and semantic ambiguity in natural language requirements," in 2018 Fifth International Symposium on Innovation in Information and Communication Technology (ISIICT). 2018;1-7.



*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*http://www.sdiarticle4.com/review-history/70231*