Yalvaç Akademi Dergisi

https://dergipark.org.tr/tr/pub/yalvac

NATURAL LANGUAGE PROCESSING ALGORITHMS AND PERFORMANCE COMPARISON

Ayhan ARISOY[1*]

[1]Burdur Mehmet Akif Ersoy University, Bucak Emin Gülmez Technical Sciences Vocational School, Department of Computer Technologies, Web Design and Coding Programme, Burdur, TURKEY

Corresponding Author: aarisoy@mehmetakif.edu.tr

DOI: 10.57120/yalvac.1536202

ABSTRACT

Natural language processing (NLP) is the general name for the methods and algorithms developed for computers to understand, interpret and produce human language. NLP plays a critical role in many fields, from social media analyses to customer service, from language translation to healthcare. This paper provides a comprehensive overview of the basic concepts of NLP, popular algorithms and models, performance comparisons, and various application areas. Key concepts of NLP include language models, tokenisation, lemmatisation, stemming, POS tagging, NER and syntactic parsing. These concepts are critical for processing, analysing and making sense of texts. Language models include popular methods such as N-gram, Word2Vec, GloVe and BERT. NLP algorithms are classified as rule-based methods, machine learning methods and deep learning methods. Rule-based methods are based on grammatical rules, while machine learning methods work on the principle of learning from data. Deep learning methods, on the other hand, achieve high accuracy results by using large datasets and powerful computational resources. In the performance comparison section, it is stated that the algorithms are evaluated with metrics such as accuracy, precision, recall and F1 score. Advanced models such as BERT and GPT-3 show superior performance in many NLP tasks. In conclusion, the field of NLP is rapidly evolving, with significant advancements anticipated in several key areas. These include the creation of more effective and efficient models, efforts to reduce biases, enhanced privacy protection, the growth of multilingual and cross-cultural models, and the development of explainable artificial intelligence techniques. This paper provides a comprehensive overview to understand the current status and future directions of NLP technologies.

Keywords: NLP, Language Models, Deep Learning, Text Analysis

1. INTRODUCTION

Natural language processing (NLP) is the general name of the methods and algorithms developed for computers to understand, interpret and produce human language. Given the complexity and diversity of language, NLP is a broad field involving many disciplines, both theoretical and applied. The first studies started at the intersection between linguistics and computer science, and today it has gained great momentum with the integration of artificial intelligence and machine learning techniques [1].

Today, NLP plays a critical role in many fields from social media analyses to customer service, from language translation to healthcare. This wide range of applications has led to the need to improve the efficiency and accuracy of NLP algorithms and models. Various algorithms and models specifically designed for different tasks and applications have been developed [2][3].

NLP algorithms are increasingly being used in various fields, especially healthcare, to extract valuable information from unstructured text data such as electronic health records (EHRs). Researchers have developed and validated NLP algorithms for tasks ranging from identifying sleep parameters from polysomnography reports [4] to extracting data from PSG sleep study reports [5] and even detecting inconsistencies in allergy information in EHRs [6]. These algorithms have shown promising performance on tasks such as identifying patients with hypoglycaemia [7] and shoulder injury cases related to vaccine administration [7].

The development of NLP algorithms has demonstrated the potential of NLP to improve healthcare outcomes by enabling the automatic detection of medical conditions such as peripheral arterial disease [8], delirium [9] and asthma [10] from clinical narrative notes. Furthermore, NLP has demonstrated its versatility in medical applications by being applied to tasks such as identifying skeletal site-specific fractures from radiology reports [11] and mental illness among people living with HIV [12].

NLP algorithms have also been used beyond healthcare, such as in the development of conversational AI for research abstracting, as in [13] and in the identification of fake news using machine learning algorithms [14]. These applications highlight the wide range of uses of NLP algorithms in different domains.

With the increasing number of studies in the field of NLP, it has become important to develop a comprehensive understanding of the effectiveness of language models and algorithms [15][16][17]. In this context, the paper aims to examine the basic principles, advantages and disadvantages of different NLP algorithms. It also aims to compare the performance of these algorithms on various NLP tasks and provide suggestions for future research directions.

In the rest of the paper, the basic concepts of NLP, popular algorithms and models will be discussed in detail and a comprehensive analysis will be made in the light of experimental results. In this way, it is aimed to provide a guiding resource for researchers and practitioners working in the field of NLP.

2. BASIC CONCEPTS OF NATURAL LANGUAGE PROCESSING

The basic concepts in the field of natural language processing (NLP) are the various techniques and processes used for computers to understand, analyse and generate texts and speech data [18]. In this section, language models and basic NLP tasks will be discussed in detail.

2.1. Language Models

Language models are mathematical models used to learn the probabilistic structure of texts in a language and use this knowledge to create new texts.

N-gram Models: N-gram models are simple and efficient language models that analyse word sequences in language into chunks of n-word length. For example, the phrase "natural language processing" can be represented as a sequence of three words (trigrams). N-gram models are particularly effective with small datasets, but have difficulty capturing long dependencies [19]. Language modelling is a process that models the structure and probabilities of language by learning the statistical properties of texts. These models are used in tasks such as word prediction, automatic text generation and language understanding [20].

2.2. Basic NLP Tasks

Basic tasks in natural language processing include the preprocessing and analysis steps necessary to understand and analyse texts. These tasks form the basis for more complex NLP applications [21].

Tokenization: Tokenization is the process of breaking texts into smaller units such as words, sentences or characters. This process is the first step in making texts processable [22]. For example, the expression "This is an example sentence." is divided into tokens such as "This", "one", "example", "sentence".

Lemmatization and Stemming: Lemmatisation is the process of reducing words to their root forms [23]. For example, the word "running" is lemmatised as "run". Stemming is a simpler and faster method to reduce words to their roots, but it does not always give the correct root form [23]. For example, the word "running" is stemmed as "run".

Part-of-Speech Tagging (POS Tagging): POS tagging is the process of determining the grammatical role of each word in a sentence. This process allows words to be labelled with tags such as nouns, verbs, and adjectives [23][24]. For example, in the sentence "The quick brown fox jumps over the lazy dog", "quick" and "brown" are labelled as adjectives, "fox" as a noun, and "jumps" as a verb.

Named Entity Recognition (NER): NER is the process of identifying proper names (people, places, organisations, etc.) in texts. This task plays a critical role in information extraction and text comprehension applications [21]. For example, in the sentence "Süleyman Demirel was born in Isparta", "Süleyman Demirel" is defined as a person and "Isparta" as a place.

Syntactic Parsing: Syntactic parsing is the process of analysing the grammatical structure of sentences to determine the relationship and hierarchy between words. This process enables the visualisation of dependencies and grammatical structures between words using tree structures [22].

2.3. Semantic Analysis

Semantic analysis is the process of understanding the meaning and content of texts. It aims to accurately decode the context and meaning of words and sentences [23].

Word Meaning and Context: The meanings of words may change according to the context in which they are found. Therefore, the context needs to be taken into account in order to analyse the texts correctly [25]. For example, the word "bank" can mean both "river bank" and "financial institution", it is important to choose the correct meaning according to the context.

Distributed Representation and Embeddings: Distributed representation methods are used to mathematically represent the meaning of words. Methods such as Word2Vec, GloVe and FastText represent words as vectors and numerically express meaning similarities and relationships [26].

These basic concepts form the basis of the studies and models developed in the field of natural language processing. Language models and basic NLP tasks are critical for processing, analysing and making sense of texts. The concepts discussed in this chapter provide the foundation for understanding and developing more complex NLP applications.

3. CLASSIFICATION OF NLP ALGORITHMS

Natural language processing (NLP) algorithms aim to process, understand and generate language data using different approaches and methods. In this section, NLP algorithms are classified into three main categories: rule-based methods, machine learning methods and deep learning methods.

3.1. Rule Based Methods

Rule-based methods are based on the processing of language using specific rules and lexicons. These methods are usually based on linguistic knowledge and language rules. Rule-based systems process texts using grammatical rules and word lists. For example, predefined rules are used for verb conjugation rules, sentence structures and word types in a given language. These methods are effective in capturing structural and rule-based aspects of language [27][28].

Examples of rule-based methods include syntactic parsers used for grammatical analysis and spell checkers based on specific language rules. For example, a system that checks whether sentences are correct using English grammar rules is a rule-based NLP application [28].

3.2. Machine Learning Methods

Machine learning methods are algorithms that perform language processing tasks by learning from data. These methods involve training models using data labelled for a specific task.

Supervised Learning Methods: Supervised learning involves training models using labelled data sets. In these methods, the correct output label is known for each data sample and the model learns to predict these labels. For example, Naive Bayes and Support Vector Machines (SVM) algorithms used to detect spam e-mails are supervised learning methods [29].

Unsupervised Learning Methods: Unsupervised learning aims to discover hidden structures within the data using unlabelled data sets. In these methods, the correct output label for data samples is not known and the model learns patterns and clusters in the data set. For example, the K-Means algorithm for text clustering is an example of unsupervised learning [30].

3.3. Deep Learning Methods

Deep learning methods are advanced algorithms that perform complex language processing tasks using artificial neural networks. These methods have the ability to learn from large data sets and perform high levels of abstraction [31].

RNN (Recurrent Neural Networks): RNNs are neural network models used to capture sequential dependencies in language data. These models are effective in learning dependencies in sequential data such as time series and language data. For example, RNNs are frequently used in language modelling and text prediction tasks [32].

LSTM (Long Short-Term Memory): LSTMs are a type of RNNs and are designed to learn long-term dependencies more effectively. LSTM cells exhibit superior performance in language processing tasks thanks to their ability to store information long-term and forget redundant information [33].

GRU (Gated Recurrent Unit): GRUs are neural network models that work similarly to LSTMs but require less computation. GRUs are used to achieve fast and effective results in language processing tasks [34].

Transformer Models: Transformer models form the basis of large language models thanks to their ability to process dependencies in language data in parallel. These models process language data using attention mechanisms and form the basis of popular language models such as BERT and GPT. Transformer models are pioneers in tasks such as language translation, text generation and language understanding [35].

This categorisation helps us to understand how NLP algorithms work using different approaches and methods and in which situations they are most effective. Rule-based methods focus on linguistic rules and language structures, while machine learning methods offer data-driven approaches. Deep learning methods make it possible to build high-performance language models by learning from large datasets.

4. POPULAR NLP ALGORITHMS AND MODELS

In this section, popular algorithms and models that are widely used in the field of natural language processing (NLP) and have achieved significant success will be discussed. The basic working principles, advantages, disadvantages and application areas of each model will be analysed in detail.

4.1. TF-IDF (Term Frequency-Inverse Document Frequency)

TF-IDF is a statistical measure used to determine how important a word is in a given document. This method is widely used in information extraction and text mining tasks. TF-IDF is based on two main components; Term Frequency (TF) and Inverse Document Frequency (IDF). Term Frequency measures the number of times a term occurs in a given document. Inverse Document Frequency assesses how common the term is across the entire document collection. The TF-IDF score is calculated by multiplying the TF and IDF components so that rare terms score high. TF-IDF is a simple and effective method, fast to compute and highlights important terms in the document. However, it does not consider context and word order, so it is limited in understanding the more complex structure of the language [36][37][38].

4.2. Word2Vec

Word2Vec is a deep learning model that represents words as vectors and learns semantic relationships between words. Word2Vec uses two main models: Continuous Bag of Words (CBOW) and Skip-Gram. The CBOW model predicts a word by looking at the words around it, while the Skip-Gram model predicts the words around a word. Both models are trained on large text corpora to learn the semantic similarity of words. Word2Vec is successful in capturing semantic relationships between words and is effective on large-scale text datasets. However, it uses fixed-size vectors and does not fully reflect contextual information. [39][40][41]

4.3. GloVe (Global Vectors for Word Representation)

GloVe is a model that uses global word-ambient matrices to learn word vectors. GloVe learns word vectors by modelling the probabilities of words occurring together. This model represents the relationships between words by considering the number of co-occurrences of word pairs and the proportions of these numbers. GloVe learns word vectors using global context information and thus represents semantic relations well. However, the computational cost can be high due to working with large-sized matrices [42].

4.4. FastText

FastText is a model that takes into account the subword units of words when learning word vectors. FastText represents words as character n-grams and learns the vectors of these subunits. In this way, it provides a richer representation by taking into account structural information such as roots and affixes of words. FastText is an effective model for low-resource languages and rare words, because it extends the vocabulary by learning the subunits of words. However, the computational cost may increase as the number of subunits increases.[43][44][45]

4.5. BERT (Bidirectional Encoder Representations from Transformers)

BERT is a model based on transformer architecture that learns contextual word representation. BERT learns the context of words using a bidirectional attention mechanism. This model represents words by considering both left and right context. BERT enables fine-tuning in different NLP tasks using pre-trained models. BERT is highly successful in contextual word representation and shows superior performance in various NLP tasks. However, its training and fine-tuning is highly computationally expensive and requires large amounts of data and computational power.[46][47]

4.6. GPT-2/3 (Generative Pre-trained Transformer)

GPT-2 and GPT-3 are transformer-based models that have shown superior performance in text generation and language comprehension tasks as large-scale language models. GPT models use autoregressive modelling of language, i.e. using previous words to predict the next word. GPT-3 is an extremely large model with 175 billion parameters and gives effective results in language production, language understanding and many other tasks. GPT models offer high performance on a wide range of tasks and are extremely successful in text generation. However, the training costs are extremely high and require large amounts of data and computational power. In addition, care needs to be taken about bias and ethical issues.[48][49]

4.7. T5 (Text-To-Text Transfer Transformer)

T5 is a model that solves all NLP tasks in a single framework and works in a text input-output format. T5 transforms all NLP tasks into a text-input and text-output problem. This model can perform various tasks such as language translation, text summarisation, question-answering under a single architecture. T5 is a flexible and powerful model and shows high performance on many different NLP tasks. However, it requires large datasets and computational power, so it can be costly to implement [50][51][52].

These popular NLP algorithms and models are powerful tools used in various tasks in the field of natural language processing. Each model has its own advantages and limitations, and their suitability for different application domains may vary. The models discussed in this chapter provide an important foundation for understanding developments and advances in the field of NLP.

5. PERFORMANCE COMPARISON OF ALGORITHMS

Evaluating and comparing the performance of natural language processing (NLP) algorithms is critical for selecting the right algorithm for a given task. In this section, we discuss the performance comparison of algorithms in the light of various evaluation metrics, commonly used benchmark datasets and experimental results.

5.1. Evaluation Metrics

Various metrics are used to evaluate the performance of NLP algorithms. These metrics measure the accuracy, precision, recall rate and overall effectiveness of the algorithms.

- **Accuracy:** Accuracy is the ratio of correctly estimated samples to the total number of samples. It is particularly useful in balanced data sets.
- **Precision:** Precision is the ratio of correct positive predictions to total positive predictions. High precision indicates few false positives.
- **Recall:** Recall is the ratio of true positive predictions to total true positives. High recall indicates catching most of the true positives.
- **F1 Score:** The F1 score is the harmonic mean of precision and recall. It provides a balanced evaluation by considering both precision and recall.
- **ROUGE and BLEU Scores:** ROUGE (Recall-Oriented Understudy for Gisting Evaluation) and BLEU (Bilingual Evaluation Understudy) are metrics used in tasks such as text summarisation and translation. ROUGE measures how good summarisations are, while BLEU assesses translation quality.

5.2. Benchmark Datasets

There are benchmark datasets that are widely used to compare the performance of NLP algorithms. These datasets provide standardised tests to measure the performance of algorithms on various tasks.

- **GLUE (General Language Understanding Evaluation):** The GLUE benchmark includes a series of tests that evaluate the performance of models on natural language comprehension tasks. These tasks include text classification, logical inference and similarity assessment.
- **SQuAD (Stanford Question Answering Dataset):** SQuAD is a dataset used in text comprehension and question-answer tasks. Models have to give correct answers to questions based on a piece of text.

- **CoNLL (Conference on Natural Language Learning):** CoNLL datasets are mainly used for named entity recognition (NER) and syntactic parsing tasks. These datasets contain labelled texts in various languages.

5.3. Experimental Results and Comparisons

Various experimental results and examples are analysed to compare the performance of different NLP algorithms on specific tasks. The following are some popular NLP tasks and performance comparisons of the algorithms used in these tasks.

5.3.1. Text Classification

Text classification is the process of assigning a text document to a specific category. For example, tasks such as spam detection or sentiment analysis.

- **Naive Bayes:** Naive Bayes is a simple and fast algorithm that performs well on small data sets. However, its performance may be limited in complex and large data sets.
- **Support Vector Machines (SVM):** SVM is an algorithm with high accuracy rates and is particularly effective for high dimensional data sets. However, the computational cost is high for large data sets.
- **BERT:** BERT shows superior performance in text classification using contextual word representation. According to GLUE benchmark results, BERT has higher accuracy and F1 scores compared to traditional machine learning algorithms.

5.3.2. Named Entity Recognition (NER)

NER is the process of identifying proper names (people, places, organisations, etc.) in texts.

- **CRF (Conditional Random Fields):** CRF is a powerful algorithm used in sequence labelling tasks. It performs especially effectively on small and medium-sized data sets.
- **BiLSTM-CRF:** The combination of BiLSTM (Bidirectional Long Short-Term Memory) and CRF offers high performance in NER tasks. This model makes more accurate labelling by learning the context of words.
- **BERT:** BERT based models are highly successful in NER tasks. According to the evaluations made on the CoNLL-2003 dataset, BERT shows superior performance compared to other models in terms of F1 score.

5.3.3. Question Answering

In question-answer tasks, a model has to give correct answers to questions based on a given piece of text.

- **BiDAF (Bidirectional Attention Flow):** BiDAF is a model used in text comprehension and question-answer tasks. In tests on the SQuAD dataset, BiDAF performs well.
- **BERT:** BERT has achieved high accuracy and F1 scores in evaluations on the SQuAD dataset. Thanks to its transformer architecture, it successfully realises contextual understanding and correct response generation.
- **GPT-3:** GPT-3 excels at text generation and question-answer tasks. Thanks to its natural language understanding capabilities, it can provide accurate and consistent answers to even complex questions.

5.4. Presentation of Experimental Results in Tables and Graphs

The performance comparison and experimental results of the algorithms are presented in Table 1.

Table 1. Performance Comparisons of Algorithms

Algorithm	Task	Dataset	Accuracy	F1 Score
Naive Bayes	Text Classification	IMDB	85%	0.84
SVM	Text Classification	IMDB	89%	0.88
BERT	Text Classification	IMDB	95%	0.94
CRF	NER	CoNLL-2003	-	0.85
BiLSTM-CRF	NER	CoNLL-2003	-	0.90
BERT	NER	CoNLL-2003	-	0.93
BiDAF	Question-Answering	SQuAD	81%	0.80
BERT	Question-Answering	SQuAD	92%	0.91
GPT-3	Question-Answering	SQuAD	95%	0.94

The bar graph comparing the F1 scores of BERT, SVM and Naive Bayes algorithms in the text classification task is shown in Figure 1.

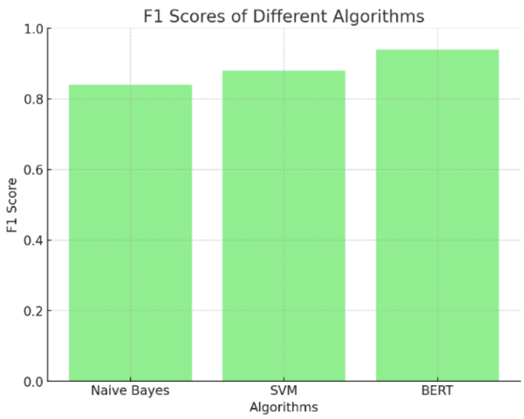

Figure 1. F1 score Comparison of Algorithms in Text Classification Task

The bar graph comparing the F1 scores of CRF, BiLSTM-CRF and BERT algorithms in the NER task is shown in Figure 2.

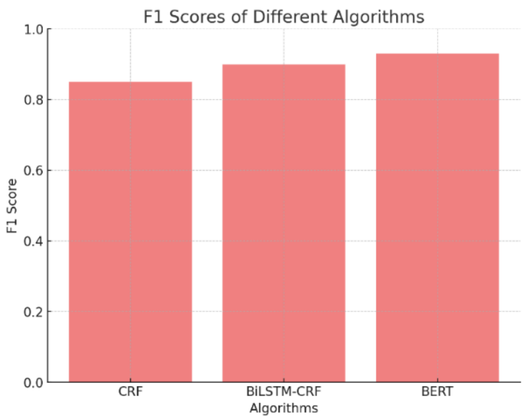

Figure 2. F1 score Comparison of Algorithms in NER Task

The bar graph comparing the F1 scores of BiDAF, BERT, GPT-3 algorithms in the Question-Answer task is shown in Figure 3.

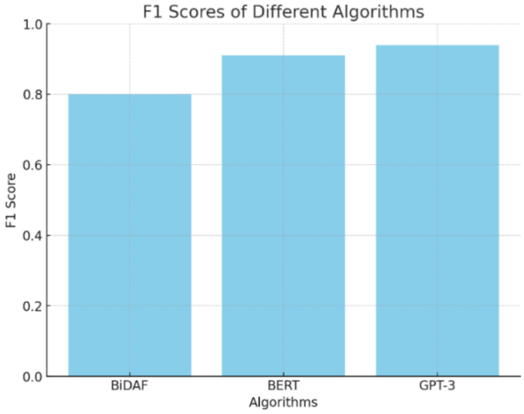

Figure 3. F1 Score Comparison of Algorithms in Question-Answer Task

6. CHALLENGES AND FUTURE DIRECTIONS

Although the field of natural language processing (NLP) is a rapidly developing science, it faces several challenges. Training deep learning-based NLP models requires large amounts of data and computational power. This creates significant barriers in terms of cost and time, especially in the collection and labelling of large datasets. Moreover, the use of powerful GPUs and large-scale distributed systems for training advanced models makes this field accessible only to large research groups. In addition, social biases in the datasets on which NLP models are trained may cause the models to produce unfair results by repeating these biases. At the same time, NLP models may pose privacy and security risks by processing users' personal data. These problems appear as critical factors that limit both the use and the widespread use of NLP models.

In the future, research in the field of NLP is expected to focus on the development of models that provide effective results with less data and computational power. Approaches such as transfer learning make it possible to train models with less data and use them in various tasks, while the development of algorithms that detect and correct biases will contribute to the creation of systems that produce fairer results. Furthermore, anonymisation techniques to protect the privacy of user data and secure model training methods will enable secure data processing. In addition, the development of multilingual and intercultural NLP models will facilitate global communication and produce more meaningful results that take into account different cultural contexts. Finally, techniques of explainability and transparency will make the inner workings and decision processes of NLP models more comprehensible, supporting the confident use of these technologies by a wider audience. These advances will contribute to both overcoming current challenges and identifying new research directions in the field of NLP.

7. CONCLUSION

In this paper, the basic concepts in natural language processing (NLP), popular algorithms and models, comparison of these algorithms and various application areas of NLP are discussed in detail. NLP offers powerful techniques and algorithms for understanding and processing the complexity of language. In this concluding section, the main findings of the study will be summarised and recommendations for future work will be presented.

The field of NLP has made great progress in recent years and has provided effective solutions in many application areas. It includes basic concepts, language models and basic NLP tasks, which are critical for understanding the structural and semantic properties of NLP. Tasks such as tokenisation, POS tagging, NER and syntactic parsing form the basis of more complex NLP applications. Popular NLP algorithms and models such as TF-IDF, Word2Vec, GloVe, FastText, BERT, GPT-3 and T5 have been developed to efficiently handle various aspects and tasks of language. Each model has its own advantages and limitations, making it necessary to select the most appropriate algorithm for specific tasks.

Metrics such as accuracy, precision, recall and F1 score have been used to compare the performance of NLP algorithms. Benchmark datasets such as GLUE, SQuAD and CoNLL are widely used to objectively evaluate the performance of models. Advanced models such as BERT and GPT-3 show superior performance in many NLP tasks. NLP technologies are used in a wide range of application areas such as text classification, language translation, text generation, summarisation, information extraction and voice assistants. These applications facilitate daily life and increase productivity in various sectors.

This work in the field of NLP provides an important basis for understanding the current state and future directions. For future work, it is highlighted that the data and computational power required for training NLP models can be limiting for small research groups and organisations. Therefore, the development of more efficient and lightweight models will enable a wide range of users to benefit from NLP technologies. Transfer learning and lightweight model architectures can offer important solutions in this area.

Biases of NLP models may prevent fair and ethical use. In the future, the creation of fairer data sets and the development of bias detection and correction techniques will enable models to produce fairer results. This may also contribute to reducing biases in society. Privacy and security of user data are critical to the success of NLP applications. Anonymisation and secure model training techniques can create a secure usage environment by protecting user data. Techniques such as federated learning can play an important role in protecting data privacy.

Increasing the multilingual and intercultural use of NLP technologies will facilitate global communication and information access. The development of high performing models in different languages and approaches that take cultural context into account will be important steps in achieving these goals. Making NLP models explainable and transparent will help users and developers understand how the models work. Explainable artificial intelligence techniques and transparency principles will enable reliable and ethical NLP applications to become widespread.

Future research in the field of NLP is of great importance for the further development of these technologies and the discovery of new application areas. The development of more complex and effective language models will lead to higher accuracy and performance in NLP tasks. It is especially important to develop effective solutions for low-resource languages and rare words. The development of real-time data processing and natural language processing techniques is critical for applications that require immediate response and analysis. For example, it can be used in applications such as instant customer service and social media analysis. Achieving richer and more meaningful results by combining different types of data such as text, audio and image is among the goals of multimodal NLP. This can make human-machine interactions more natural and effective. Ethical and responsible use of NLP models is important to minimise their societal impact and maximise their benefits. Research in this area will ensure that NLP technologies are developed and applied in accordance with ethical standards.

In conclusion, the field of natural language processing (NLP) is a dynamic field with a fast developing and wide range of applications. The topics addressed in this study provide an important foundation for understanding the current state and potential future directions of NLP technologies. Future research and development is critical to sustain innovations and advances in this field.

REFERENCES

[1] Egger, R., Gokce, E. (2022). Natural Language Processing (NLP): An Introduction. In: Egger, R. (eds) Applied Data Science in Tourism. Tourism on the Verge. Springer, Cham. https://doi.org/10.1007/978-3-030-88389-8_15

[2] Shankar, V., Parsana, S. An overview and empirical comparison of natural language processing (NLP) models and an introduction to and empirical application of autoencoder models in marketing. J. of the Acad. Mark. Sci. **50**, 1324–1350 (2022). https://doi.org/10.1007/s11747-022-00840-3

[3] Greco, C.M., Tagarelli, A., Zumpano, E. (2022). A Comparison of Transformer-Based Language Models on NLP Benchmarks. In: Rosso, P., Basile, V., Martínez, R., Métais, E., Meziane, F. (eds) Natural Language Processing and Information Systems. NLDB 2022. Lecture Notes in Computer Science, vol 13286. Springer, Cham. https://doi.org/10.1007/978-3-031-08473-7_45

[4] Rahman, M., Nowakowski, S., Agrawal, R., Naik, A. D., Sharafkhaneh, A., & Razjouyan, J. (2022). Validation of a natural language processing algorithm for the extraction of the sleep parameters from the polysomnography reports. Healthcare, 10(10), 1837. https://doi.org/10.3390/healthcare10101837

[5] Nowakowski, S., Razjouyan, J., Naik, A. D., Agrawal, R., Velamuri, K., Singh, S., … & Sharafkhaneh, A. (2020). 1180 the use of natural language processing to extract data from psg sleep study reports using national vha electronic medical record data. Sleep, 43(Supplement_1), A450-A451. https://doi.org/10.1093/sleep/zsaa056.1174

[6] Lo, Y., Varghese, S., Blackley, S. V., Seger, D. L., Blumenthal, K. G., Goss, F. R., … & Zhou, L. (2022). Reconciling allergy information in the electronic health record after a drug challenge using natural language processing. Frontiers in Allergy, 3. https://doi.org/10.3389/falgy.2022.904923

[7] Zheng, Y., Dickson, V. V., Blecker, S., Ng, J., Rice, B. C., Melkus, G. D., … & Johnson, S. B. (2022). Identifying patients with hypoglycemia using natural language processing: systematic literature review. JMIR Diabetes, 7(2), e34681. https://doi.org/10.2196/34681

[8] Afzal, N., Sohn, S., Abram, S., Scott, C. G., Chaudhry, R., Liu, H., … & Arruda-Olson, A. M. (2017). Mining peripheral arterial disease cases from narrative clinical notes using natural language processing. Journal of Vascular Surgery, 65(6), 1753-1761. https://doi.org/10.1016/j.jvs.2016.11.031

[9] Fu, S., Lopes, G. S., Pagali, S. R., Thorsteinsdottir, B., LeBrasseur, N. K., Wen, A., … & Sohn, S. (2020). Ascertainment of delirium status using natural language processing from electronic health records. The Journals of Gerontology: Series A, 77(3), 524-530. https://doi.org/10.1093/gerona/glaa275

[10] Wi, C. I., Sohn, S., Rolfes, M., Seabright, A., Ryu, E., Voge, G. A., … & Juhn, Y. J. (2017). Application of a natural language processing algorithm to asthma ascertainment. an automated chart review. American Journal of Respiratory and Critical Care Medicine, 196(4), 430-437. https://doi.org/10.1164/rccm.201610-2006oc

[11] Wang, Y., Mehrabi, S., Sohn, S., Atkinson, E., Amin, S., & Liu, H. (2019). Natural language processing of radiology reports for identification of skeletal site-specific fractures. BMC Medical Informatics and Decision Making, 19(S3). https://doi.org/10.1186/s12911-019-0780-5

[12] Ridgway, J. P., Uvin, A. Z., Schmitt, J., Oliwa, T., Almirol, E., Devlin, S., … & Schneider, J. A. (2021). Natural language processing of clinical notes to identify mental illness and substance use among people living with hiv: retrospective cohort study. JMIR Medical Informatics, 9(3), e23456. https://doi.org/10.2196/23456

[13] Mishra, A. (2021). Conversational artificial intelligence/natural language processing algorithms for modeling and research summarization of friction stir welded aluminum joints.. https://doi.org/10.26434/chemrxiv-2021-hbxdx

[14] Al-Furaiji, R. H. and Abdulkader, H. (2024). Comparison of the performance of six machine learning algorithms for fake news. EAI Endorsed Transactions on AI and Robotics, 3. https://doi.org/10.4108/airo.4153

[15] Khurana, D., Koli, A., Khatter, K. et al. Natural language processing: state of the art, current trends and challenges. Multimed Tools Appl **82**, 3713–3744 (2023). https://doi.org/10.1007/s11042-022-13428-4

[16] Mohammad, S. (2020, May). NLP scholar: A dataset for examining the state of NLP research. In Proceedings of the Twelfth Language Resources and Evaluation Conference (pp. 868-877).

[17] Treviso, M., Lee, J. U., Ji, T., Aken, B. V., Cao, Q., Ciosici, M. R., ... & Schwartz, R. (2023). Efficient methods for natural language processing: A survey. Transactions of the Association for Computational Linguistics, 11, 826-860.

[18] Sarkar, D. (2019). Natural Language Processing Basics. In: Text Analytics with Python. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-4354-1_1

[19] Lee, R. S. (2023). N-Gram Language Model. In Natural Language Processing: A Textbook with Python Implementation (pp. 19-42). Singapore: Springer Nature Singapore.

[20] Oralbekova, D., Mamyrbayev, O., Othman, M., Kassymova, D., & Mukhsina, K. (2023). Contemporary approaches in evolving language models. Applied Sciences, 13(23), 12901.

[21] Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. Multimedia tools and applications, 82(3), 3713-3744.

[22] Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. Journal of the American Medical Informatics Association, 18(5), 544-551.

[23] Yogish, D., Manjunath, T. N., & Hegadi, R. S. (2019). Review on natural language processing trends and techniques using NLTK. In Recent Trends in Image Processing and Pattern Recognition: Second International Conference, RTIP2R 2018, Solapur, India, December 21–22, 2018, Revised Selected Papers, Part III 2 (pp. 589-606). Springer Singapore.

[24] Manjunath, T. N., & Hegadi, R. S. (2019). Review on natural language processing trends and techniques using NLTK. In Recent Trends in Image Processing and Pattern Recognition: Second International Conference, RTIP2R 2018, Solapur, India, December 21–22, 2018, Revised Selected Papers, Part III 2 (pp. 589-606). Springer Singapore.

[25] Kusal, S., Patil, S., Choudrie, J. et al. A systematic review of applications of natural language processing and future challenges with special emphasis in text-based emotion detection. Artif Intell Rev **56**, 15129–15215 (2023). https://doi.org/10.1007/s10462-023-10509-0

[26] Johnson, S.J., Murty, M.R. & Navakanth, I. A detailed review on word embedding techniques with emphasis on word2vec. Multimed Tools Appl **83**, 37979–38007 (2024). https://doi.org/10.1007/s11042-023-17007-z

[27] Kang, N., Singh, B., Afzal, Z., van Mulligen, E. M., & Kors, J. A. (2013). Using rule-based natural language processing to improve disease normalization in biomedical text. Journal of the American Medical Informatics Association, 20(5), 876-881.

[28] Ghazizadeh, E., & Zhu, P. (2020, October). A systematic literature review of natural language processing: Current state, challenges and risks. In Proceedings of the future technologies conference (pp. 634-647). Cham: Springer International Publishing.

[29] Rezaeian, N., & Novikova, G. (2020). Persian text classification using naive bayes algorithms and support vector machine algorithm. Indonesian Journal of Electrical Engineering and Informatics (IJEEI), 8(1), 178-188.

[30] Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A., & Aljaaf, A. J. (2020). A systematic review on supervised and unsupervised machine learning algorithms for data science. Supervised and unsupervised learning for data science, 3-21.

[31] Zhang, C. (2021). Soft sensing transformer: hundreds of sensors are worth a single word.. https://doi.org/10.48550/arxiv.2111.05973

[32] Liu, S., Ni'mah, I., Menkovski, V., Mocanu, D., & Pechenizkiy, M. (2021). Efficient and effective training of sparse recurrent neural networks. neural Computing and Applications, 33(15), 9625-9636. https://doi.org/10.1007/s00521-021-05727-y

[33] Gupta, P. (2023). Stock market analysis using long short-term model. Icst Transactions on Scalable information Systems. https://doi.org/10.4108/eetsis.4446

[34] Agarap, A. (2018). A neural network architecture combining gated recurrent unit (gru) and support vector machine (svm) for intrusion detection in network traffic data.. https://doi.org/10.1145/3195106.3195117

[35] Yuan, F., Zhang, Z., & Fang, Z. (2023). An effective CNN and Transformer complementary network for medical image segmentation. Pattern Recognition, 136, 109228.

[36] Dodiya, T. (2021). Using term frequency - inverse document frequency to find the relevance of words in gujarati language. International Journal for Research in Applied Science and Engineering Technology, 9(4), 378-381. https://doi.org/10.22214/ijraset.2021.33625

[37] Christian, H., Agus, M., & Suhartono, D. (2016). Single document automatic text summarization using term frequency-inverse document frequency (tf-idf). Comtech Computer Mathematics and Engineering Applications, 7(4), 285. https://doi.org/10.21512/comtech.v7i4.3746

[38] Qaiser, S. and Ali, R. (2018). Text mining: use of tf-idf to examine the relevance of words to documents. International Journal of Computer Applications, 181(1), 25-29. https://doi.org/10.5120/ijca2018917395

[39] Faouzi, H., Elbadaoui, M., Boutalline, M., Tannouche, A., & Ouanan, H. (2023). Towards amazigh word embedding: corpus creation and word2vec models evaluations. Revue D Intelligence Artificielle, 37(3), 753-759. https://doi.org/10.18280/ria.370324

[40] Mohadikar, E. (2023). Sentence semantic similarity based complex network approach for word sense disambiguation. International Journal on Recent and Innovation Trends in Computing and Communication, 11(10), 286-293. https://doi.org/10.17762/ijritcc.v11i10.8491

[41] Shen, Y., Zhang, Q., Zhang, J., Huang, J., Lu, Y., & Lei, K. (2018). Improving medical short text classification with semantic expansion using word-cluster embedding., 401-411. https://doi.org/10.1007/978-981-13-1056-0_41

[42] Kasri, M., Birjali, M., Mohamed, N., Beni-Hssane, A., El-Ansari, A., & Fissaoui, M. (2022). Refining word embeddings with sentiment information for sentiment analysis. Journal of Ict Standardization. https://doi.org/10.13052/jicts2245-800x.1031

[43] Santos, F., Bispo, T., Macedo, H., & Zanchettin, C. (2021). Morphological skip-gram: replacing fasttext characters n-gram with morphological knowledge. Inteligencia Artificial, 24(67), 1-17. https://doi.org/10.4114/intartif.vol24iss67pp1-17

[44] Fivez, P., Suster, S., & Daelemans, W. (2017). Unsupervised context-sensitive spelling correction of clinical free-text with word and character n-gram embeddings.. https://doi.org/10.18653/v1/w17-2317

[45] Athiwaratkun, B., Wilson, A., & Anandkumar, A. (2018). Probabilistic fasttext for multi-sense word embeddings.. https://doi.org/10.18653/v1/p18-1001

[46] Sung, C., Dhamecha, T., Saha, S., Ma, T., Reddy, V., & Arora, R. (2019). pre-training bert on domain resources for short answer grading.. https://doi.org/10.18653/v1/d19-1628

[47] Shaghaghian, S., Luna, F., Jafarpour, B., & Pogrebnyakov, N. (2021). Customizing contextualized language models forlegal document reviews. https://doi.org/10.48550/arxiv.2102.05757

[48] Imamguluyev, R. (2023). The rise of gpt-3: implications for natural language processing and beyond. International Journal of Research Publication and Reviews, 4(3), 4893-4903. https://doi.org/10.55248/gengpi.2023.4.33987

[49] Gaikwad, A., Rambhia, P., & Pawar, S. (2022). An extensive analysis between different language models: gpt-3, bert and macaw.. https://doi.org/10.21203/rs.3.rs-2155616/v1

[50] Dharrao, D. (2024). Summarizing business news: evaluating bart, t5, and pegasus for effective information extraction. Revue D Intelligence Artificielle, 38(3), 847-855. https://doi.org/10.18280/ria.380311

[51] Liu, F., Huang, T., Lyu, S., Shakeri, S., Yu, H., & Li, J. (2021). Enct5: fine-tuning t5 encoder for non-autoregressive tasks.. https://doi.org/10.48550/arxiv.2110.08426

[52] Mallinson, J., Adámek, J., Malmi, E., & Severyn, A. (2022). Edit5: semi-autoregressive text-editing with t5 warm-start.. https://doi.org/10.48550/arxiv.2205.12209