CHAPTER «PHILOLOGICAL SCIENCES»

ANALYSIS OF THE THEORETICAL FOUNDATIONS OF NEURAL NETWORK MODELING OF LANGUAGE UNIT RECOGNITION

Oleksii Dovhan[1]

DOI: https://doi.org/10.30525/978-9934-26-310-1-9

Abstract. In the context of linguistic science and the integration of the mathematical paradigm into humanitarian discourse, the analysis and processing of natural language (in particular, neural network modeling of language categories) is an important and urgent task. Research by a number of authors and a series of experiments show that artificial neural networks of various types and kinds with different parameterizations can significantly optimize linguistic research: accelerate, deepen, integrate into various scientific fields, etc. At the same time, the use of artificial neural networks in linguistics is an important area of work, as well as a powerful and productive tool for a number of relevant studies, which, however, requires careful analysis and development of implementation strategies. *The purpose* of the article is to analyze the features of neural network modeling of language units recognition as an effective method of cognition within the anthropocentric paradigm of research. The solution of such research tasks determines the logic of presentation of the studied material in the article: introduction, systematization of achievements in the theory and practice of modeling as a universal tool of scientific cognition in general and theoretical justification of neural network modeling in the context of linguistic paradigm. *Methodology* of the study is based of the method of analyzing scientific research was updated, which led to the search and analysis of scientific publications related to neural network modeling (in particular, language units). We analyzed more than 60 recent scientific studies and publications covering aspects of the problem under study,

[1] PhD in Philology, Doctoral Student at the Department of Slavic Languages, Drahomanov Ukrainian State University, Ukrainе

which we evaluated based on their relevance, methodological specificity, and scientific novelty. Thus, the research methods outlined in this article allowed us to conduct a thorough analysis of the state and prospects for the development of the theoretical foundations of neural network modeling of language unit recognition. The analysis of the latest scientific research and publications has made it possible to determine the role and place of artificial neural networks of various types and their specifications in the process of modeling language units. The above made it possible to identify the main trends in working with text data aimed at improving the quality of their processing, generation, etc. Results of the survey showed that one of the core tasks of modern linguistic science is to understand the language polysystem, the peculiarities of its structure and the nature of its functioning. In addition, the discursive nature of language practices in the context of data interconnection is also important for documenting language structures and verbal practices that are socially determined. That is why the use of neural networks as a tool for conducting local linguistic and integrated scientific research involving the mathematical paradigm is gaining popularity. Neural network modeling of language categories in this context is the basic basis for such research. *Practical implications.* The actualization of neural network modeling of linguistic categories is not only one of the means of studying linguistic polysystems, but also an objective criterion of checking the truth of linguistic knowledge. *Value/ originality.* Neural network modeling of language unit recognition has a significant impact on the development, improvement, and evolution of modern linguistic research. The effectiveness of neural network modeling of language categories creates opportunities for a deeper understanding of the studied linguistic objects, phenomena and processes, encouraging linguists to develop various linguistic models that could solve practical linguistic problems (information retrieval, machine translation, natural language processing, knowledge extraction and localization from text, etc.).

1. Introduction

The rapid development of information technologies has significantly changed our everyday life, approaches to data search, processing, and analysis. At the same time, information technologies have had a two-sided effect: by influencing a number of processes in society, they have also changed themselves, their quality, nature and structure, adapting to social demand.

First of all, we are talking about the emergence of intelligent data processing systems, as well as Deep Learning and Machine Learning methods (support vector machine, decision tree, naive Bayesian classifier, etc.). Another important milestone was the creation of neural network models that are structurally and functionally similar to the neural network of the human brain [15].

At the same time, the functioning of information of any nature naturally had a number of features, the core of which was the recursiveness of its existence, the essence of which is the need to analyze the current data segment through the localization of its previous links [48]. This recursiveness also means the representation of the research subject in an inseparable connection with it, i.e., the situation when we describe such an object by actualizing its fractal (likening the part to the whole) properties [16].

The above characteristics of the language polysystem are organically actualized in the process of integrating the tools of other sciences into the humanitarian (in general) and linguistic (in particular) paradigms. We are talking about the approach to a language polysystem as a data set, which is consistent with the use of Data Science tools [54], as well as in the context of the interconnection of cultural and ecological diversity data for documenting language structures and verbal practices [24]. This approach is representative for understanding the linguistic and pragmatic features of modeling the representation of linguistic units and mastering the network of relationships between the neural network and the senses of textual data in the context of computer, mathematical and corpus linguistics, etc.

For our research, the Data Science approach is fundamental, since it is about analyzing a language polysystem as a data set with its own characteristics, specifics, and peculiarities of use. This approach to analyzing this type of data is relevant because the collection, processing and processing of a large amount of information, the frequency of contact with it, as well as its analytical and synthetic processing are crucial in the problem of using the latter (the pragmatic component of everyday life).

The above issue is related to the special interest of researchers and the public in the problems of data processing in the context of the pragmatic foundations of this process. For example, the development of voice and online assistants, a number of chatbots, etc. creates the need for more and more thorough actualization (mastering lexical, semantic, lexicographic layers, etc.) of the language polysystem. In its turn, the above-mentioned leads to the increasingly intensive development of Natural Language Processing, an integrated scientific field that combines mathematical and computer linguistics, programming, mathematical statistics, computer analysis and synthesis of natural language [4].

A number of computer programs actively use Natural Language Processing (data retrieval, text summarization, categorization, vectorization (transformation of lexical units (words) into vectors representing the input text), aggregation (merging) of information content, etc. However, the problem with Natural Language Processing is that a language polysystem is a complex construct that is functionally an element of a social contract. We are talking about the magical, pragmatic, and other functions of a language polysystem, in particular, that it acts as a form and environment for the existence of the intended sense (expressed or written). In turn, the aforementioned construct has a number of properties, among which the following are particularly noteworthy:

1. Discreteness – fleeting, intermittent interpretability of semantics, which is characterized by situationality, emphasis on certain categories or properties, etc.

2. Categoricality is represented through structuredness (we are talking about categorical symbols that are encoded as signals for communication through various channels: audio, graphic (in this case, gestural and direct writing) or correlation (we are talking about cohesion or coherence, semantic continuity and predictability, which allows us to distinguish between a particular thought by formally presenting it in sentences, paragraphs, etc.).

3. Symbolism is graphically represented or conditional data that has a meaning, is based on a social contract, and produces a certain response. Here, the symbol acts not just as a conventional designation of a certain subject, phenomenon or object, but as an epistemological marker that denotes the ontological nature of the language polysystem. Thus, symbolism allows us to position the language polysystem as an entity with a unique coding system, data transmission channels, etc., characterized by stability and reliability of existence [42].

2. Neural network modeling as a tool for analyzing language units

The problem of the theoretical foundations of neural network modeling of language category recognition is integrated and multifaceted, which is why a number of scientists are actively developing it, researching it in different contexts. In this regard, we consider the work of Cîmpeanu I. A. [11], which investigates the relevance of using artificial neural networks (in particular, artificial recurrent neural networks (RNN)).

In his research, the author illustrates the relevance of using artificial recurrent neural networks (RNNs) through their use in two applications in different fields of activity. In addition, the researcher emphasizes a number of peculiarities of using neural networks: he notes the need for human intervention in their work to obtain an optimal result, lists the limitations of such entities, etc.

Pan X., Wang Y., Qi Y. [44] continue to research artificial neural networks in their work. In the analyzed research, the authors position the artificial neural network as an information processing system created by modeling the structure and logical thinking of the human brain. The researchers see the uniqueness of such a network in its nonlinearity and ability to perform data processing tasks similar to the human brain.

In particular, the scientists note that the above tools are ideal for working with both structured (relational) and unstructured (non-relational) data, which allows processing information that is difficult to identify by conventional methods. In their work, the researchers present the origin, specifics of types and development of artificial neural networks, and summarize the progress of applied research in natural language processing (NLP).

The work of artificial neural networks in the context of the modeling process is presented in the work of Pagkalos M., Chavlis S., Poirazi P. [43]. In their research, the authors consider the method of computational modeling as an indispensable tool for visualizing data. In particular, the researchers study the role of dendritic computing in network-level operations, which has long remained virtually unexplored.

The above is due to the fact that existing tools do not allow generating productive and realistic network models that take into account dendrites. The researchers note that spiking neural networks (SNNs) are effective in this case, but an important drawback of the latter is their simplification (ignoring a number of essential properties of the objects of research). At the same time, the use of models with morphologically detailed neuronal models is costly, which makes it impossible to update them in the operation of large networks.

The peculiarities of the existence of neural network models and the specifics of working with them are investigated in Guest O., Martin A. E. [22]. In their research, the authors reflect on what can be an indicator of the absolute performance of a particular model in certain scientific fields.

The purpose of the researchers' reasoning is to generalize certain parametric data about a particular model to statements about their biological counterpart (the brain and its abilities), as well as the neurocognitive abilities of the above systems. Scientists note that conclusions about the above features are often based on the specifics of the model's performance of a particular task on the scale of approximation (replacing a complex object with a simple one as close as possible to the original) of human behavior and brain activity in general.

The specifics of the existence of data in their cells, as well as the peculiarities of working with them in the context of natural language processing (NLP) are presented in Baruah R. D., Organero M. M. [7]. In particular, in their research, the authors focused on the peculiarities of the development and progress of sensory, communication and computer technologies in the context of data cells.

The researchers emphasize that the existence, extraction and analysis of data in the above-mentioned data centers have a number of features. The main one is that data in such cells or environments can be obtained not only from the sources or objects of monitoring identified by the authors, but also directly from the environment or discourse in which the information is used.

Features of data grouping in the process of natural language processing (NLP) are presented in the work of Loureiro M. V., Derby S., Wijaya T. K. [35]. The researchers emphasize that the peculiarities of expressiveness, lexical richness, etc. of the analyzed language polysystem directly affect the amount of computing resources. This specificity of the process of modeling language categories allowed the authors to propose a new approach to cluster modeling based on conceptual entities. According to the scientists, an entity is a linguistic and diagnostic representation of the concepts of ontological reality, which contains a number of relational (ordered) data.

The above approach allowed the authors to extract vector representations of entities from (I) the encyclopedic corpus (using a language model) and (II) the knowledge base (using an artificial graph neural network (GNN)). The researchers claim that this approach outperformed other modern thematic models in all parameters (metrics) of connectivity, as well as in data extraction performance. Thus, explicit data encoded in graph structures represent a higher percentage of textual coherence than implicit data encoded in contextualized structures of language models.

Methodological features of word embeddings in modern Machine Learning models are presented in Umer M., Imtiaz Z., Ahmad M., Nappi M., Medaglia C., Choi G. S., Mehmood A. [59]. In their research, the researchers analyze the features of updating the artificial convolutional neural network (CNN) model for text data classification through a series of experiments. The proposed neural network model was trained by generating labeled Fast Text words on publicly available datasets, six benchmark datasets (including Ag News, Amazon Full and Polarity, Yahoo Question Answer, Yelp Full and Polarity).

One of the main tasks of natural language processing (NLP) is the classification of text data, which became the object of Magalhães D., Lima R. H., Pozo A. [37]. This research presents the application of an evolutionary grammar-based approach to the design of deep neural networks (DNNs) using models based on convolutional neural networks of long short-term memory (LSTM) and graph neural networks (GNNs). The authors propose different grammars that are designed to take into account the peculiarities of each type of network, as well as a number of specifications that will help to verify their impact on the created designs and the performance of the generated neural network models.

Amjad M., Gelbukh A., Voronkov I., Saenko A. [3] continue to research the features of text data classification using various artificial neural networks in their work. The research demonstrates the features of working with an artificial convolutional neural network (CNN), which does not require semantic or syntactic data and is productive at the level of lexical units – words. The researchers also emphasize the performance of updating an artificial recurrent neural network (RNN), which is able to efficiently classify text data, outperforming other neural networks in classifying sequences in text data.

The issue of actualization of Deep Learning in the context of efficient use of hardware in the process of natural language processing (NLP) is presented in the work of Selvam K. P., Brorsson M. [51]. In their research, the authors analyze the parameterization of the output characteristics that affect the overall balance of the data processing system, when the model receives the optimal amount of resources. The authors claim that their proposed Deep Learning Inference Performance Prediction Model (DIPPM) predicts inference latency, energy consumption, and memory updates for a given input Deep Learning model on an NVIDIA A100 GPU.

The specifics of approaches to the problem of Machine Learning are highlighted in Fu Y., Wang S., Li X., Li D., Li Y., Liao J., Zheng J. [17]. Thus, according to the authors, this type of training is productive, built on the basis of an artificial recurrent neural network (RNN), which is actually used for the modeling process. In their research, the researchers developed a Machine Learning approach based on the architecture of a recurrent neural network (RNN) to work with information by creating a certain filter for the input data. Using the above filter, the authors managed to significantly improve the process of data analysis by the neural network model and, as a result, obtain more accurate results of its work based on the ATLAS experiment.

The peculiarities of the process of semantic learning of a neurobiologically limited model are presented in the work of Henningsen-Schomers M. R., Garagnani M., Pulvermüller F. [25]. In their research, the authors present the peculiarities of working with the above model in the context of learning specific and abstract concepts by a neural network model, both with and without verbal notation. The researchers note that the above process was based on the mechanisms of Hebbian learning (an approach to Machine Learning based on training a neural network with many layers, its specificity is the passage of data through successive layers of neurons, each of which calculates the output values that become the beginning of the process in the next layer).

The above is related to the specifics of the existence of self-normalizing artificial neural networks (SNN), which became the object of the work of Madasu A., Rao V. A. [36]. In their research, the authors argue that feed-forward artificial neural networks (FNNs) are a productive tool for performing a number of Machine Learning tasks. Scientists analyze the features of artificial neural networks, emphasizing that their goal is to obtain a normalized output for a normalized input in the process of natural language processing (NLP).

Natural language processing (NLP) is an important component of the use of artificial neural networks, which is why the work of Hafeez R., Anwar M. W., Jamal M. H., Fatima T., Espinosa J. C. M., López L. A. D., Ashraf I. [23] on the above issue in the context of Machine translation is relevant to our research. In this paper, the authors present the peculiarities of building an information retrieval system, emphasizing that normalization and morphological analysis are the core of this process. The researchers also emphasize that in order to build a correct information search, it is necessary to identify the correct root in different words, which will make it possible to create a morphological analyzer.

The research of sign language processing as a component of natural language analysis (NLA) and various approaches to processing its data was carried out by Dias T. S., Junior J. J. A. M., Pichorim S. F. [13]. The authors present in their research a comparison of gesture recognition models based on a tool glove with a manual approach to feature extraction. For manual feature extraction (classical pattern recognition method, and without it), the method of direct sequential feature selection (SFS) was updated, which ensured the selection of the most relevant indicators and the analysis of various classifiers. For automatic feature extraction, the authors used artificial recurrent neural networks with long-term short-term memory (LSTM), Gated Recurrent Unit (GRU) models, and, as an additional model, the Ensemble GRU model. The structure of the latter was stratified into three equal parts: initial, gesture period, and final period.

Sharma S., Saxena V. P. [53], who in their work develop the issue of an automatic sign language recognition system. For this purpose, the researchers used an artificial convolutional neural network (CNN), which, in their opinion, surpasses other types of networks by the ability to use a wide range of visual tasks. The authors note that improving the results in a number of indicators (accuracy, evaluation, speed, etc.) requires a methodological restructuring and a significant increase in the resource base. The researchers believe that the way out of the above situation is to use an artificial convolutional neural network (CNN) with equal-scale layers and various filters.

Damaneh M. M., Mohanna F., Jafari P. [12] continue to research the above problem, revealing such an aspect of it as the structure of an artificial neural network Deep Learning for identifying the static movement of a hand gesture in sign language. The proposed structure includes an artificial convolutional neural network (CNN) and a classical non-intelligent feature extraction method. In the context of the framework proposed by the authors, the hand gesture image (after some processing and background removal) goes through a three-stage feature extraction process, which will significantly improve the quality of data processing. The peculiarity of the methodology is that although the process has a triune nature, the processing of each of these areas is carried out in parallel with the specific selection of specific features.

Another aspect of the research of natural language processing (NLP) in the context of human-computer interaction is presented in the work of Sethia D., Singh P., Mohapatra B. [52]. In their research, the authors continue to work on sign language recognition in the context of modern trends in image processing, in particular, the use of artificial neural networks in the process of their interpretation. For this purpose, they commonly used an artificial convolutional neural network (CNN), which allowed them to optimize the recognition process, since hand gestures differ from one person to another, becoming a problem with a non-linear (ambiguous) nature.

The problem of data preparation in the context of natural language processing (NLP) is addressed by Schriner J. [50]. In his research, the author presents data mining and Machine Learning software for sequence-to-sequence models in the field of computational linguistics. In his work, the author discusses in detail the features of sequence-to-sequence learning with the actualization of code and project results (for example, Esperanto pronunciation prediction, etc.).

The features of natural language processing (NLP) in the context of speech recognition are presented in Liu J., Gan J., Chen K., Wu D., Pan W. [34]. In this research, the authors note that speech recognition technology is a popular area of research in the field of artificial intelligence, especially in the context of Deep Learning. According to the authors, most modern speech recognition models are productive for major language polysystems, but they are not fruitful for working with isolated languages and in conditions of resource scarcity. The authors believe that the way out of this situation is to use a Deep Learning approach with the actualization of a simple and effective model for recognizing isolated language polysystems or minority language polysystems.

The research of the features of Deep Learning in the context of speech polysystem recognition is continued in Pardede H. F., Adhi P., Zilvan V., Ramdan A., Krisnandi D. [45]. The research notes that the interest in updating Deep Learning in the context of text data processing is related to the ability of these technologies to model language categories, track pronunciation difficulties, recognize syntax features and language rules, etc. In the analyzed research, we present Deep functions based on artificial convolutional neural networks (CNNs) for the Indonesian language with a large continuous recognition vocabulary. The main feature of the training model presented by the authors is its discriminative nature (in contrast to the usual organization of such models, whose training is built generatively).

Continuation of the analysis of natural language processing (NLP) using artificial neural networks in the context of language polysystem recognition is the subject of research by Adolfi F., Bowers J. S., Poeppel D. [1]. In the analyzed work, the authors combined the experiments on the above issues into a single system for processing, analyzing and synthesizing text data to evaluate the performance of state-of-the-art artificial neural networks. The above allowed the researchers to find out the peculiarities of correlations between speech manipulations in text and natural speech; to present the conditions under which artificial neural networks represented out-of-distribution stability, reproducing the peculiarities of human perception; to identify the optimal conditions under which the predictions of neural network models and humans differ, etc.

Reza S., Ferreira M. C., Machado J. J. M., Tavares J. M. R. [49] continue to research the features of speech recognition in the context of natural language processing (NLP). In their work, the authors argue that text data recognition is a complex and integrated problem, as well as the possibilities of applying the results of the above studies. Scientists emphasize that the use of artificial neural networks for natural language analysis (NLA) is relevant, which is why the combination of different types of neural network models in this work is productive. For instance, researchers claim that elements of residual neural networks (ResNet) before recurrent cells of artificial neural networks (RNN) improve the accuracy of the result and significantly reduce the training time.

Systems for automatic recognition of a language polysystem with a large vocabulary and other natural language processing (NLP) programs cannot work without generating appropriate language models, the specifics of which are presented in Mukhamadiyev A., Mukhiddinov M., Khujayarov I., Ochilov M., Cho J. [39]. In the analyzed research, the authors emphasize that the vast majority of the above pre-trained language models are focused on common language polysystems (English, German, French, etc.). The researchers note that there are no publicly available language datasets for less common languages, one of which is Uzbek. In turn, this allowed them to conclude that it is important to research and create language models for languages with limited resources.

The continuation of the research of the problem of natural language processing (NLP) and the peculiarities of the presentation of its categories is presented in the work of Veličković P. [60]. In his research, the author positions graphs as the main form of representation of data obtained by humans from biological systems, which is explained by the universal nature of graph structures. Thus, in the author's opinion, the above-mentioned formations are related to the peculiarities of the structure of patterns (objects from which copies or sets of repeated objects are made) that we observe in natural and artificial systems. The researcher argues that it is productive to apply Machine Learning to the processing of graphical, textual, etc. data of a language polysystem that can be represented on the basis of graph representations, which, in turn, will produce integration of scientific fields, contributing to their development.

Wu L., Chen Y., Shen K., Guo X., Gao H., Li S., Long B. [63] continue to analyze the use of graph representation in Machine Learning in their work. In their research, the authors presented a comprehensive overview of artificial graph neural networks (GNNs) used for natural language processing (NLP). In particular, the researchers proposed a new taxonomy (classification features) of graph neural networks for natural language processing (NLP): they grouped all available studies into three areas (graph construction, features of their representation, and graph-based encoder-decoder models).

Further research of the features of data processing in the context of their visual representation (visualization) is presented in the work of

Cheng C. H., Ji Z. T. [10]. In their research, the authors focused on the specifics of natural language processing (NLP) for the text data they collected. The researchers present the peculiarities of their application of the Dirichlet latent distribution method for modeling topics and obtaining their categorization based on two coherence metrics (tools for assessing the quality of topic models) to assign a class to each article. The authors emphasize that in order to obtain an optimal classification result, edge weights need a representative. Thus, according to the scientists, it is advisable to use the Neo4j tool (to display nodes and edge weights to obtain the final data) for the above-mentioned visualization (visual representation).

The specifics of natural language processing (NLP) in the context of perceiving handwritten input of Tibetan characters are analyzed in the work of Zhang G., Wang W., Zhang C., Zhao P., Zhang M. [64]. In their research, the authors emphasize that the above process is quite ambiguous due to the specifics of writing certain hieroglyphs, as well as the morphological features of similar characters. The researchers point out that the recognition accuracy can be improved by using deep neural networks (DNNs), which are the best compromise between the accuracy of the results obtained with their help and the speed of the analysis process.

The authors emphasize that it is possible to reduce the parameterization of artificial neural networks being trained and maintain acceptable accuracy by using a neural network model called HUTNet. The aforementioned neural network model is based on the internal relationship between the number of floating point operations per second (FLOPs) performed and a number of system requirements for memory access.

A research of the functioning of artificial neural networks for performing tasks close to human ones is described in Kanwisher N., Khosla M., Dobs K. [29]. In their research, the authors present the properties and functions of the biological nervous system, correlating them with the "behavior" of artificial neural networks optimized for certain tasks. The paper analyzes the possibility of artificial neural networks reflecting the behavioral and neural characteristics of a human performing the same task. Against the background of the actualization of the above-mentioned approach to the functioning of biological neural networks, the researchers explain the peculiarities of the visual and auditory systems in the context of the behavioral and neural levels, as well as the possibility of their integration into artificial neural networks.

The analysis of neural representations of visual perception is devoted to the work of Lee J., Jung M., Lustig N., Lee J. H. [31]. In the above research, the authors analyze ten handwritten digits and six visual objects using an artificial convolutional neural network (CNN) and a person undergoing magnetic resonance imaging. The researchers note that after tuning the artificial convolutional neural network model using the pre-trained VGG16 model to recognize visual stimuli from the categories of numbers and objects, they conducted a representational similarity analysis (RSA) using neural activations in a human undergoing magnetic resonance imaging and the representations of the convolutional neural network model features in all 16 classes. The authors emphasize that the coded representation of the convolutional neural network model reflected the hierarchical topography of the human visual system. Thus, the mapping of features in the lower convolutional layers (Conv) was similar to neural representations in the early visual areas and parietal cortex, including the posterior cingulate cortex.

Wang J., Xie H., Wang F. L., Lee L. K., Wei M. [61] continue to study the peculiarities of the use of graph convolutional neural networks (GNN) and work with them. In their research, the authors assert the effectiveness of using artificial graph neural networks (GNNs) in working with sense. The scientists note that existing models for analyzing complex schemes of transitions from one element to another build a local and global graph of sense. At the same time, according to the researchers, reuse on the local graph is low-frequency, and working with neighboring vertices throughout the learning process is quite difficult due to computational complexity and limited space. To solve the above problems, the authors propose to use an artificial graph neural network (GNN), which can be used to build intra- and inter-session dependencies of elements for senses recommendations. Scientists note that it is advisable to build a local graph of sessions taking into account repetitions, as this will allow encoding internal dependencies between elements and generating a representation of senses based on positional awareness.

Features of working with heterogeneous graph networks are presented in Li J., Liu M., Wang Y., Zhang D., Qin B. [32]. The researchers note that parsing (automated collection and structuring of information using software or a specific service) of discourse for a multilateral dialogue is a tool for identifying the structure of the above environment and connections in dialogic speech and obtaining a graph of discourse dependencies.

The authors emphasize that a number of existing models designed to solve the above problem prove the influence of data on the speaker, but none of them provides data for its representation in the context of analyzing the discourse structure of dialogues. In their research, the researchers propose their own HG-MDP model, which is based on a heterogeneous graph neural network (GNN) to encode the above-mentioned dialog graphs. In addition, the aforementioned artificial neural network is capable of iterative (repeated) updating for aggregation (the process of combining data from different sources to provide a more thorough information representation) of speaker vertices and utterances.

The effectiveness of Deep Learning in natural language processing (NLP) was studied in the work of Klemen M., Krsnik L., Robnik-Šikonja M. [30]. In their research, the authors argue that the use of Deep Learning in natural language processing (NLP) is more productive than other methods. This is due to the fact that this type of learning has a special ability to extract certain features, properties, etc. with the establishment of a number of correlations between these data. The authors emphasize that the most productive neural architectures are long-term short-term memory (LSTM) and the neural network model-transformer (BERT), which are appropriate for use on large pre-trained language models (for example, the aforementioned BERT). It should be noted that most modern methods of natural language processing (NLP) are focused on English, while the analysis of less resource-intensive language polysystems is not sufficiently developed, which, by the way, directly confirms the relevance of the research under consideration.

The peculiarities of applying Deep Learning methods for natural language processing (NLP) on the examples of classification and text generation are described in Tuffery S. [58]. In his research, the author focuses on artificial recurrent neural networks (RNNs), as well as transformer models (in particular, BERT). The researcher has classified texts using an artificial recurrent neural network (RNN) and the transformer model DistilBERT. In addition, the scientist presented a number of examples of the implementation of artificial deep neural networks (DNNs) (in particular, artificial recurrent neural networks (RNNs)).

The researcher notes that in the case of a long source text or simple vocabulary, the ratio between the total number and the number of different words will be high, allowing for the generation of a new text. The author emphasizes that textual data can be structured in such a way that artificial convolutional neural networks (CNNs) can be naturally applied to it.

Iparraguirre-Villanueva O., Guevara-Ponce V., Ruiz-Alvarado D., Beltozar-Clemente S., Sierra-Liñan F., Zapata-Paulini J., Cabanillas-Carbonell M. [27] continue to research the features of working with the above-mentioned long-term short-term memory (LSTM), a type of artificial recurrent neural network (RNN). The researchers note that the aforementioned type of artificial neural network is based on sequence and is actively used in the processes of analysis, generation, and forecasting in working with text data due to its ability to analyze long-term dependencies. The authors emphasize that their research combines an LSTM-network and a sifting technique to generate text from a corpus as input, and creates a model to find the most productive way to analyze and extract words from context.

The features of automatic textual data summarization are the basis of the work by Ghadimi A., Beigy H. [18]. In their research, the authors note that in addition to researching the specifics of sentence construction or feature learning and finding solutions for textual data summarization, it is productive to actualize Deep Learning. The researchers see the solution to the above issues in the use of a submodular graph convolutional summarizer (SGCS), a method of extractive summarization of multi-document data in which the data source name (DSN) guarantees minimal performance.

In their work, the researchers consider the specifics of the data source name (DSN), in particular, the shortcomings and improvements in its operation that allow obtaining generalized data, in terms of support for sentence construction and feature learning based on graph structure. The research updates a formal approach that allows us to represent the role and structure of the above-mentioned updates of the submodular graph convolutional summarizer (SGCS) in the context of guaranteeing minimum performance.

The specifics of neural architecture search (NAS) and the effectiveness of automatic training of neural network models are highlighted in Li Y., Cao R., He Q., Xiao T., Zhu J. [33]. In their research, the authors argue that most neural architecture systems are unreliable due to a gap in their structure. In their research, the authors attempt to improve the performance and stability of neural architectures by minimizing the gaps between their representations. The researchers note that the core of this process is a general representation of the abbreviation for modeling Neural Architecture Search with Distributed Architecture Representations (ArchDAR).

According to the authors, the best result is achieved through joint learning, which ensures the integration of distributed representations with leading architecture search methods. The core of the ArchDAR implementation is a differentiated model of architecture search and testing of learned architectures on language category modeling (based on Penn Treebank). According to the authors, this approach provides greater stability, which leads to faster structural convergence when working with a neural network model with a differentiated architecture.

The core problem of natural language processing (NLP) in the work of Mishra B. K., Jain S. [38] call the ambiguity of semantics, and the researchers see the solution to this issue in the use of a Word Sense Disambiguation. With the help of this tool, it is possible to determine the appropriate meaning of polysemous words in an arbitrary context using computational methods. Naturally, any language polysystem is inherent in ambiguity, which can be analyzed using a variety of knowledge-based tools, supervised and unsupervised approaches. The researchers argue that the last ten years in India have seen the launch of a number of digital services, the functioning of which is naturally based on natural language processing (in particular, Hindi or other Indian languages). It is worth noting that due to limited resources for natural language processing (recognition of polysemous words, etc.), digital services in India use the simplest version of neural network technology (with the possibility of integration with IndoWordNet).

Bölücü N., Can B., Artuner H. [8] discuss the peculiarities of sense representation as a way of expressing the meaning of a text that can be processed by a machine to perform a certain task in the context of natural language processing (NLP). In their work, the researchers present a model of semantic parsing based on artificial neural networks that helps to obtain a semantic representation of a given sentence. The authors emphasize that they used the semantic representation of each sentence to create semantically rich sentences, which allowed them to perform an external evaluation of the proposed semantic analyzer.

The aforementioned semantic analyzer is based on a self-attention mechanism (reflexive mechanism). The latter allows analyzing semantic relations between words in a sentence and building a semantic representation in the structure of the Universal Conceptual Cognitive Annotation (UCCA) semantic annotation. In its turn, the above annotation is by its nature a cross-linguistic graph-based semantic representation that allows for sentiment analysis as well.

The aforementioned sentiment analysis is one of the important components in natural language processing (NLP): it is primarily about determining the degree of expressiveness of the text under research, which was the object of research in the work of Mutinda J., Mwangi W., Okeyo G. [40]. The researchers emphasize that there are two groups of text representation methods: lexicon-based and Machine Learning. At the same time, both groups of methods have their limitations: for example, pre-trained word embedders (Word2Vec, Glove, bidirectional encoding from transformers (BERT), etc.) perform vectorization, which takes into account the following parameters: distance between words, similarity, etc.

The sentiment classification model (LeBERT) presented in the analyzed research integrates the processing of sentiment vocabulary, n-grams, a feedforward neural network model (BERT), and an artificial convolutional neural network (CNN). In the model created by the researchers, the aforementioned tools were updated to carry out the process of vectorizing words selected from a fragment of the input text data. The artificial convolutional neural network (CNN) is used as a deep neural network classifier to represent the features and obtain the sentiment class at the output.

Further work on sentiment detection in text data is presented in Gu T., Zhao H., He Z., Li M., Ying D. [21]. The research is devoted to aspectual sentiment analysis, the essence of which is to analyze the polarity of sentiment in a certain ranked range. The most productive tool for the above analysis, according to the researchers, is an artificial graph convolutional neural network (GCN).

At the same time, the authors emphasize that most existing studies focus on extracting word-aspect-word contextual dependencies and dependency trees based on the sentence itself without updating a significant amount of external data related to textual data (in particular, discourse). The researchers emphasize that the problem of intelligent word capture beyond grammatical distance and boundary markers does not allow to update the functionality of artificial graph convolutional neural networks (GCN).

The above allowed the researchers to propose their own version of an artificial graph convolutional neural network (GCN) that combines external data (sentiment vocabulary and morphological data) (EK-GCN). Thus, the authors conducted a statistical research of the parts and built their matrix (to fully take into account the influence of negative words, degrees of comparison, etc. that represent emotions, classify them, etc.).

The specifics of natural language processing (NLP) in the context of sentiment analysis are presented in Suhartono D., Purwandari K., Jeremy N. H., Philip S., Arisaputra P., Parmonangan I. H. [55]. In their research, the authors propose a Deep Learning-based approach to sentiment recognition based on drug recall data from the UCI Machine Learning repository. The researchers position the proposed approach as an alternative to Deep Learning models with an architecture that incorporates cells with artificial convolutional neural networks (CNNs).

The aforementioned integration schemes (Word2vec, GloVe word embedding) were empirically evaluated for their predictive performance in the architecture of an artificial convolutional neural network (CNN). Based on the comparison of the Deep Learning architecture with RoBERTa, it is clear that the architecture of the aforementioned neural network model-transformer (BERT) outperforms both in training and validation. At the same time, the artificial convolutional neural network (CNN) models, in which the Glove word embedding is updated, showed better results in the testing process.

The above-mentioned direction of data processing in the context of their fusion was studied in the work of Jaafar N., Lachiri Z. [28]. In their research, the researchers analyze the phenomenon of multimodal fusion, which is one of the most controversial areas in affective computing performed using artificial neural networks. The core complexity of the above process is the combination of data from different sources and modalities, especially in the context of identifying certain emotional reactions. The essence of the problem of identifying certain emotions lies in their ambiguous nature, as it is a differentiated phenomenon that is difficult to describe with a number of algorithms and classifiers.

In their research, the authors present an approach to solving the above problem based on four methods of multimodal fusion: audiographic, videographic, textual, and additional information with Deep Learning methods to process the stream of integrated data. The features of natural language processing (NLP) in the context of analyzing input audio data are presented in the work of Girirajan S., Pandian A. [19]. In their research, the authors emphasize the relevance of processing input audio data in the context of converting noisy input speech into an improved output signal. The researchers argue that the problem of speech enhancement is relevant in the context of natural language processing (NLP), in particular, automatic speech recognition (ASR), as well as mobile voice communication systems.

The authors emphasize that most of the research in the field of speech enhancement is conducted for more widespread languages (English, Chinese, etc.), while Indian regional languages remain under-researched. In this paper, the researchers propose a two-stage architecture for improving Tamil speech analysis and processing based on a convolutional neural network (CNN). The latter performs real-time enhancement of the input audio data in a single channel or audio track, etc.

The problem of large language models (LLM), in particular the OpenAI tool "Chat Generative Pre-Trained Transformer", better known as "ChatGPT", is devoted to the work of Alberts I. L., Mercolli L., Pyka T., Prenosil G., Shi K., Rominger A., Afshar-Oromieh A. [2]. In their research, the authors focused on the features of large language models based on the aforementioned artificial multilayer recurrent neural networks (RNNs), which are trained on a large amount of data (large corpora of texts) and result in the creation of a human-like text.

The researchers emphasize a number of differences between traditional language models and the aforementioned big language models. For instance, the main feature, according to them, is a number of methods for predicting the logic of word ordering in a sentence. In the case of traditional language models, statistical methods are used to calculate the probability of actualization of a particular lexical item. The principle of operation of large language models (in particular, the aforementioned ChatGPT) is that this prediction is made by using transformer-based models that parallelize the processing process. This nature of the aforementioned large-scale language model allows it to create (or, more precisely, compile and summarize data) computer code, fiction, journalistic and other texts.

The specifics of the use of transformer models in the context of natural language processing (in particular, automatic translation) are presented in the work of Badawi S. [6]. In his research, the author describes the specifics of the transformer model functioning, in particular, he presents the mechanism of attention and the dynamics of its performance. The researcher argues that the proposed model has proven itself in the automatic translation of texts into English, French, and German with large resources.

The scientist claims that the author's model is the first-ever transformation-based neural machine translation model for the Kurdish language that uses lexical items that share a common set of specifications across the entire dataset. Updating the above neural network language model was made possible by combining all available parallel corpora of the Kurdish and English language polysystems. In turn, this allowed us to create a large corpus, which became the material for developing the transformer model.

Further research of working with large language models is presented in the work of Perrine P. [47]. In his research, the author emphasizes the inextricable link between large language models and Machine Learning, as well as the features of their scaling and Deep Learning in the context of the continuation of empirical linguistic methods. At the same time, the author contrasts this group of methods with others based on rules limited by a nativist perspective.

The scientist emphasizes that the inaccessibility of modern master's programs with limited resources for researchers due to a number of circumstances produces the emergence of nativist biases among the latter. That is why the researcher emphasizes the need to publish large open-source language models so that empirical and hybrid approaches are accessible and updated by researchers.

Natural language processing (NLP) in the context of knowledge or data tracking was studied by Ni Q., Wei T., Zhao J., He L., Zheng C. [41]. The authors note that knowledge tracking is an essential component of the online education system (its relevance has naturally increased in our country during the COVID-19 pandemic and the war) along with the rapid development of adaptive online learning. The researchers emphasize that the deep knowledge tracking model has a number of features, the main one being excessive parameterization. The aforementioned specificity leads to the inability to fully distinguish the features of the searched data due to the heterogeneity of their structure and small volume.

The authors propose to solve the above problems in knowledge tracking by changing the approach to the searched data. Thus, the researchers want to base the search on heterogeneous hierarchical differentiation (HHSKT). The essence of the proposed method is that hierarchical heterogeneous knowledge structures and short-term memory improvements will be actualized to build models of the impact of various sequences of interaction with people.

The research of artificial neural networks in the context of a dialogic recommender system (CRS) based on natural language processing (NLP) is presented in Wu J., Yang B., Li D., Deng L. [62]. In their research, the authors analyze the work of end-to-end dialogic recommender systems (CRS), a subtype that simulates the task of recommendations and monologue simultaneously, paralleling the processing of all text data streams. At the same time, end-to-end dialog recommender systems include a knowledge graph and a transformer, which allows them to perform natural language processing (NLP) more efficiently.

Natural language processing (NLP) includes many different aspects, one of which is the reliability of data that exists in a particular discourse or environment. In Dixit D. K., Bhagat A., Dangi D. [14], the authors analyze the peculiarities of solving the problem of fake (false) data in the context of neural network modeling based on supervised and unsupervised learning methods. The researchers emphasize that the aforementioned studies do not provide accurate and measurable results. According to the scientists, this is due to various reasons, the main ones being the imbalance of the data sets used to train the neural network model, the low efficiency of parameterization of such a model, and the irrelevant choice of the studied features, etc.

To solve a number of the above problems, the researchers propose to use a fundamentally new approach to detecting fake (false) data. Thus, according to the authors, it is advisable to start by obtaining information from the ISOT dataset with its subsequent cleaning. Next, it is necessary to perform preliminary processing, which has a three-component structure: stemming (the process of separating the root part of a word from its ending and other affixes), removal of stop words (words that do not have a special sense load) and tokenization (the process of dividing text data using letters, spaces, punctuation marks and numbers into words, phrases, sentences, paragraphs, etc. [56]).

Scientists emphasize that after pre-processing the data, it is important to graduate the features in the process of extracting them: they talk about the need to update various features, giving preference to those based on recognizing the essence of names. Moreover, features of short dimensionality should be selected using a modified ensemble model of independent component analysis. In addition, the researchers emphasize that it is advisable to update the artificial hybrid convolutional neural network (CNN) based on the Levy algorithm, which is based on the honey badger flight algorithm, which allows detecting fake (false) data.

The authors also note that a series of experiments were modeled using Python software with different performance metrics (accuracy – 95%, specificity – 97%, sensitivity – 98%) and allowed to test the performance of the proposed method. The material for the experimental work was the ISOT dataset, on which the proposed method showed better detection results and higher selection accuracy.

Bozkir A. S., Dalgic F. S., Dalgic F. C., Aydos M. [9] continue to research the peculiarities of working with data in the context of their truthfulness. In the analyzed research, the authors focused on phishing (a type of online fraud in which attackers, posing as representatives of reliable institutions, lure confidential information from people: usernames, passwords and credit card data) as a component of the above problem and the specifics of its detection using artificial neural networks with minimal input data costs.

In their paper, the researchers propose a new artificial deep neural network (DNN) model for identifying phishing URLs, which they call GramBeddings. According to the scientists, the above model has a number of features: updating n-gram tabs, which are fast and do not require preliminary training; eliminating the need for data at the level of lexical units and their sublevels, implementing a smart and efficient n-gram selection mechanism, and using the attention mechanism. The researchers note that the proposed artificial neural network contains a controlled and automated n-gram selection and filtering mechanism, as well as a new neural network architecture that combines a four-channel information flow through cascading layers of artificial convolutional neural network (CNN), long short-term memory (LSTM) recurrent neural networks (RNN), and attention.

The authors emphasize that such an actualization structure contributed to the detection of discriminative multi-level character patterns without additional (manual) intervention and the ability to facilitate prediction. Thus, according to the scientists, the neural network they propose provides end-to-end and high-performance real-time output, prediction based on language diagnostics, and eliminates the need for any services or manual functionality. The authors claim that experiments have shown the effectiveness of their proposed approach (the accuracy of the new artificial deep neural network model is 98.7%, which is higher than the results of the corresponding models). Moreover, the scientists note that a comparative research conducted on several datasets showed the benefits of using GramBeddings.

The above leads to the conclusion that the emergence of artificial neural networks has become an organic continuation of data science in general and developments in the field of artificial intelligence in particular. The first neural network models appeared back in the 50s of the last century, but due to limited technical capabilities, they were simple and low-functional. A real breakthrough in this area occurred with the development of technological progress, in particular, thanks to the technologies for rapid training and updating of such models. We are talking about the emergence of graphics processors, whose power allowed for parallel processing of large data sets (in our case, corpora of texts).

3. Methods

This article analyzes the theoretical foundations of neural network modeling of language unit recognition in the context of integrating the mathematical paradigm into humanities (linguistic) research, which has led to the interdisciplinarity and universalization of the issues under consideration.

In the course of the research, the method of analyzing scientific research was updated, which led to the search and analysis of scientific publications related to neural network modeling (in particular, language units). We analyzed more than 60 recent scientific studies and publications covering aspects of the problem under research, which we evaluated based on their relevance, methodological specificity, and scientific novelty.

Thus, the research methods outlined in this article allowed us to conduct a thorough analysis of the state and prospects for the development of the theoretical foundations of neural network modeling of language unit recognition. The analysis of the latest scientific research and publications has made it possible to determine the role and place of artificial neural networks of various types and their specifications in the process of modeling language units. The above made it possible to identify the main trends in working with textual data aimed at improving the quality of their processing, generation, etc.

The results of the research can be useful for developing a strategy for working with text data, in particular in the context of updating the most productive tools for working with them. We are talking about the selection of tools in accordance with the purpose of a particular linguistic research: classification, generation, processing, vectorization, etc.

4. Results

The research found that today artificial neural networks are a powerful tool for processing, generating and analyzing: text data (Bing, ChatCPT, Frase.io, LeiaPix, Merlin, Notion.AI, Humata, Jasper AI, Rytr.me, Writeme. AI, Writesonic, etc.); audio data (Adobe Speech Enhancer, Audio AI, Humtap, Narakeet, Podcastle, Sounfraw, Wisecut, etc.); graphic data (Adobe Firefly, Colourlab AI, Dalle 2, Dalle mini, Leonardo. AI, Luma AI, MidJourney, Peebley, Playground, Stable Diffusion, Tome, Waymark, Masterpiece, etc.); video data (2short.AI, Captions, DaVinci Resolve, Descript, D-ID, InVideo, Kaiber AI, Neurodub, Picture, Runway, SpiritMe, Synthesia, Topaz Video AI, Vidyo AI, etc.). The relevance of neural network modeling is related to the fact that it is a tool for solving complex problems that cannot be solved using traditional programming methods.

The results of the research show that one of the core tasks of modern linguistic science is to understand the language polysystem, the peculiarities of its structure and the nature of its functioning.

In addition, the discursive nature of language practices in the context of data interconnection is also important for documenting language structures and verbal practices that are socially determined. That is why the use of neural networks as a tool for conducting local linguistic and integrated scientific research involving the mathematical paradigm is gaining popularity. Neural network modeling of language categories in this context is the basic basis for such research.

The main idea of using neural network models for analyzing, recognizing, generating, verifying, vectorizing, etc. language categories is to find implicit correlations between textual data. In this case, neural networks become a tool for detecting patterns in this type of data with the ability to build promising models of language category development [20]. Thus, the most popular way to update neural network models in linguistics is to use them for classification of language categories (classification by meaning, grammatical classification, classification of morphological forms, etc.), generation, vectorization, stemming, stop word removal, tokenization, etc.

5. Discussion

As mentioned above, a neural network is a static model inspired by the structure and functioning of neurons in the human brain. The main specificity of their functioning is the interconnectedness of each element, which facilitates and speeds up data processing and allows them to be processed in parallel. The theoretical foundations of neural network modeling are based on Machine Learning algorithms, whose task is to search for correlations (in this case, interconnections and patterns).

This process takes place by adjusting the weights of connections between artificial neurons, with the degree of adaptability of these weights being set by the learning algorithm. The simplest option for building a neural network is a fully connected neural network – a perceptron, which has input, hidden, and output layers. Note that each layer can contain one or more neurons: their number depends on the degree of abstraction of the problem under research and, as a result, the number of hidden layers, the number of copies between them, etc. increases.

The most common neural network models for linguistic research are feed-forward neural networks, recurrent neural networks, and convolutional neural networks. The above types of neural network models are the most productive for recognizing language categories, because they are able to distinguish complex linguistic structures, etc. This ability of neural network models is due to the fact that their aforementioned training algorithm constantly scans the input data, adjusting the weights of the connections to obtain the optimal result.

6. Conclusions

The research found that the theoretical foundations of neural network modeling of language unit recognition have a significant impact on the development, improvement and evolution of modern linguistic research. Thus, the aforementioned problem is integrated and multifaceted, as represented by the work of Fu Y., Wang S., Li X., Li D., Li Y., Liao J., Zheng J. [17], which outlines an approach to Machine Learning based on the architecture of an artificial recurrent neural network (RNN), which allowed researchers to significantly improve the process of data analysis by neural network models.

Summarizing, we can draw the following conclusions. In the context of linguistic science and the integration of the mathematical paradigm into humanitarian discourse, the analysis and processing of natural language (in particular, neural network modeling of language categories) is an important and urgent task. Research by a number of authors and a series of experiments show that artificial neural networks of various types and kinds with different parameterizations can significantly optimize linguistic research: accelerate, deepen, integrate into various scientific fields, etc. At the same time, the use of artificial neural networks in linguistics is an important area of work, as well as a powerful and productive tool for a number of relevant studies, which, however, requires careful analysis and development of implementation strategies.

At the same time, the theoretical foundations of neural network modeling of language category recognition are an urgent scientific linguistic problem, the research of which is of an integrated nature, since its results will be useful for linguistics in general, in particular, Translation Theory, Environmental Linguistics, Computational Linguistics, Mathematical Linguistics Machine and Deep Learning, Mathematical Statistics, Data Science, Programming, etc.

References:

1. Adolfi F., Bowers J. S., Poeppel D. (2023). Successes and critical failures of neural networks in capturing human-like speech recognition. *Neural Networks*, 162, 199–211. DOI: https://doi.org/10.1016/j.neunet.2023.02.032

2. Alberts I. L., Mercolli L., Pyka T., Prenosil G., Shi K., Rominger A., Afshar-Oromieh A. (2023). Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be? *European journal of nuclear medicine and molecular imaging*, 1–4. DOI: https://doi.org/10.1007/s00259-023-06172-w

3. Amjad M., Gelbukh A., Voronkov I., Saenko A. (2023). Comparison of Text Classificatio n Methods Using Deep Learning Neural Networks. In *Computational Linguistics and Intelligent Text Processing: 20th International Conference, CICLing 2019, La Rochelle, France, April 7–13, 2019, Revised Selected Papers, Part II* (pp. 438–450). Cham: Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-24340-0_33

4. Arivazhagan N., Van Vleck T. T. (2023). Natural language processing basics. *Clinical Journal of the American Society of Nephrology*, 18(3), 400–401. DOI: https://doi.org/10.2215/CJN.0000000000000081

5. Arivazhagan, N., & Van Vleck, T. T. (2023). Natural language processing basics. *Clinical Journal of the American Society of Nephrology*, 18(3), 400–401.

6. Badawi S. (2023). A Transformer-Based Neural Network Machine Translation Model for the Kurdish Sorani Dialect. *UHD Journal of Science and Technology*, 7(1), 15–21. DOI: https://doi.org/10.21928/uhdjst.v7n1y2023.pp15-21

7. Baruah R. D., Organero M. M. (2023). Explicit Context Integrated Recurrent Neural Network for Sensor Data Applications. *arXiv preprint arXiv:2301.05031*. DOI: https://doi.org/10.48550/arXiv.2301.05031

8. Bölücü N., Can B., Artuner H. (2023). A Siamese Neural Network for Learning Semantically-Informed Sentence Embeddings. *Expert Systems with Applications*, 214, 119103. DOI: https://doi.org/10.1016/j.eswa.2022.119103

9. Bozkir A. S., Dalgic F. C., Aydos M. (2023). GramBeddings: A New Neural Network for URL Based Identification of Phishing Web Pages Through N-gram Embeddings. *Computers & Security*, 124: 102964. DOI: https://doi.org/10.1016/j.cose.2022.102964

10. Cheng C. H., Ji Z. T. (2023). A weighted-link graph neural network for lung cancer knowledge classification. *Applied Intelligence*, 1–19. DOI: https://doi.org/10.1007/s10489-022-04437-9

11. Cîmpeanu I. A. (2023). Applications that Use Recurrent Neural Networks and Their Impact on People's Lives. In *Education, Research and Business Technologies: Proceedings of 21st International Conference on Informatics in Economy (IE 2022)* (pp. 199–208). Singapore: Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-19-6755-9_16

12. Damaneh M. M., Mohanna F., Jafari P. (2023). Static hand gesture recognition in sign language based on convolutional neural network with feature extraction method using ORB descriptor and Gabor filter. *Expert Systems with Applications*, 211, 118559. DOI: https://doi.org/10.1016/j.eswa.2022.118559

13. Dias T. S., Junior J. J. A. M., Pichorim S. F. (2023). Comparison between handcraft feature extraction and methods based on Recurrent Neural Network

models for gesture recognition by instrumented gloves: A case for Brazilian Sign Language Alphabet. *Biomedical Signal Processing and Control*, 80, 104201. DOI: https://doi.org/10.1016/j.bspc.2022.104201

14. Dixit D. K., Bhagat A., Dangi D. (2023). An accurate fake news detection approach based on a Levy flight honey badger optimized convolutional neural network model. *Concurrency and Computation: Practice and Experience*, e7382. DOI: https://doi.org/10.1002/cpe.7382

15. Ekpenyong M. E., Udoh I. I. (Eds.). (2023). *Current Issues in Descriptive Linguistics and Digital Humanities: A Festschrift in Honor of Professor Eno-Abasi Essien Urua*. Springer Nature. URL: https://cutt.ly/57F1Wsb (date of application: 18.05.23).

16. Fractal. *Cambridge Dictionary*: web site. Available at: https://cutt.ly/i7OmAac (date of application: 18.05.23).

17. Fu Y., Wang S., Li X., Li D., Li Y., Liao J., Zheng J. (2023). Hierarchical neural network: Integrate divide-and-conquer and unified approach for argument unit recognition and classification. *Information Sciences*, 624, 796–810. DOI: https://doi.org/10.1016/j.ins.2022.12.050

18. Ghadimi A., Beigy H. (2023). SGCSumm: An extractive multi-document summarization method based on pre-trained language model, submodularity, and graph convolutional neural networks. *Expert Systems with Applications*, 215, 119308. DOI: https://doi.org/10.1016/j.eswa.2022.119308

19. Girirajan S., Pandian A. (2023). Real-Time Speech Enhancement Based on Convolutional Recurrent Neural Network. *Intelligent Automation & Soft Computing*, 35.2. DOI: https://doi.org/10.32604/iasc.2023.028090

20. Greengard S. (2023). Computational Linguistics Finds its Voice. *Communications of the ACM*, 66(2), 18–20. DOI: https://doi.org/10.1145/3575666

21. Gu T., Zhao H., He Z., Li M., Ying D. (2023). Integrating external knowledge into aspect-based sentiment analysis using graph neural network. *Knowledge-Based Systems*, 259, 110025. DOI: https://doi.org/10.1016/j.knosys.2022.110025

22. Guest O., Martin A. E. (2023). On logical inference over brains, behaviour, and artificial neural networks. *Computational Brain & Behavior*, 1–15. DOI: https://doi.org/10.1007/s42113-022-00166-x

23. Hafeez R., Anwar M. W., Jamal M. H., Fatima T., Espinosa J. C. M., López L. A. D., Ashraf I. (2023). Contextual Urdu Lemmatization Using Recurrent Neural Network Models. *Mathematics*, 11(2), 435. DOI: https://doi.org/10.3390/math11020435

24. Harrison K. D. (2023). Environmental Linguistics. *Annual Review of Linguistics*, 9. DOI: https://doi.org/10.1146/annurev-linguistics-031220-013152

25. Henningsen-Schomers M. R., Garagnani M., Pulvermüller F. (2023). Influence of language on perception and concept formation in a brain-constrained deep neural network model. *Philosophical Transactions of the Royal Society* B, 378(1870), 20210373. DOI: https://doi.org/10.1098/rstb.2021.0373

26. Huang M. et al. (2020) Sentiment strength detection with a context-dependent lexicon-based convolutional neural network. *Information Sciences*, 520, 389–399.

27. Iparraguirre-Villanueva O., Guevara-Ponce V., Ruiz-Alvarado D., Beltozar-Clemente S., Sierra-Liñan F., Zapata-Paulini J., Cabanillas-Carbonell M. (2023). Text prediction recurrent neural networks using long short-term memory-dropout. *Indonesian Journal of Electrical Engineering and Computer Science*, 29(3), 1758–1768. DOI: https://doi.org/10.11591/ijeecs.v29.i3.pp1758-1768

28. Jaafar N., Lachiri Z. (2023). Multimodal fusion methods with deep neural networks and meta-information for aggression detection in surveillance. *Expert Systems with Applications*, 211, 118523. DOI: https://doi.org/10.1016/j.eswa.2022.118523

29. Kanwisher N., Khosla M., Dobs K. (2023). Using artificial neural networks to ask 'why'questions of minds and brains. *Trends in Neurosciences*. DOI: https://doi.org/10.1016/j.tins.2022.12.008

30. Klemen M., Krsnik L., Robnik-Šikonja M. (2023). Enhancing deep neural networks with morphological information. *Natural Language Engineering*, 29(2), 360–385. DOI: https://doi.org/10.1017/S1351324922000080

31. Lee J., Jung M., Lustig N., Lee J. H. (2023). Neural representations of the perception of handwritten digits and visual objects from a convolutional neural network compared to humans. *Human Brain Mapping*. DOI: https://doi.org/10.1002/hbm.26189

32. Li J., Liu M., Wang Y., Zhang D., Qin B. (2023). A speaker-aware multiparty dialogue discourse parser with heterogeneous graph neural network. *Cognitive Systems Research*. DOI: https://doi.org/10.1016/j.cogsys.2023.01.001

33. Li Y., Cao R., He Q., Xiao T., Zhu J. (2023). Learning Reliable Neural Networks with Distributed Architecture Representations. *ACM Transactions on Asian and Low-Resource Language Information Processing*. DOI: https://doi.org/10.1145/3578709

34. Liu J., Gan J., Chen K., Wu D., Pan W. (2023). Recognition Method of Wa Language Isolated Words Based on Convolutional Neural Network. In *Machine Learning for Cyber Security: 4th International Conference, ML4CS 2022*, Guangzhou, China, December 2–4, 2022, Proceedings, Part III (pp. 631–640). Cham: Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-20102-8_49

35. Loureiro M. V., Derby S., Wijaya T. K. (2023). Topics as Entity Clusters: Entity-based Topics from Language Models and Graph Neural Networks. *arXiv preprint arXiv:2301.02458*. DOI: https://doi.org/10.48550/arXiv.2301.02458

36. Madasu A., Rao V. A. (2023). Effectiveness of self normalizing neural networks for text classification. In *Computational Linguistics and Intelligent Text Processing: 20th International Conference, CICLing 2019, La Rochelle, France, April 7–13, 2019, Revised Selected Papers, Part II* (pp. 412–423). Cham: Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-24340-0_31

37. Magalhães D., Lima R. H., Pozo A. (2023). Creating deep neural networks for text classification tasks using grammar genetic programming. Applied Soft Computing, 110009. DOI: https://doi.org/10.1016/j.asoc.2023.110009

38. Mishra B. K., Jain S. (2023). Word Sense Disambiguation for Hindi Language Using Neural Network. In *Advancements in Smart Computing and Information Security: First International Conference, ASCIS 2022*, Rajkot, India,

November 24–26, 2022, Revised Selected Papers, Part I (pp. 14–25). Cham: Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-23092-9_2

39. Mukhamadiyev A., Mukhiddinov M., Khujayarov I., Ochilov M., Cho J. (2023). Development of Language Models for Continuous Uzbek Speech Recognition System. *Sensors*, 23(3), 1145. DOI: https://doi.org/10.3390/s23031145

40. Mutinda J., Mwangi W., Okeyo G. (2023). Sentiment analysis of text reviews using lexicon-enhanced bert embedding (LeBERT) model with convolutional neural network. *Applied Sciences*, 13(3), 1445. DOI: https://doi.org/10.3390/app13031445

41. Ni Q., Wei T., Zhao J., He L., Zheng C. (2023). HHSKT: A learner-question interactions based heterogeneous graph neural network model for knowledge tracing. *Expert Systems with Applications*, 215, 119334. DOI: https://doi.org/10.1016/j.eswa.2022.119334

42. Onyshchenko K., Daniiel Ya., Kameniev R. (2020) Analiz metodiv obrobky pryrodnoi movy [Analysis of natural language processing methods]. *Informatsiini systemy ta tekhnolohii*: materialy 9-yi Mizhnar. nauk.-tekhn. konf., 17-20 lystopada 2020 r. Kharkiv: Drukarnia Madryd, 2020. S. 186–190. Available at: https://cutt.ly/v7PjIMs (date of application: 18.05.23). (in Ukrainian)

43. Pagkalos M., Chavlis S., Poirazi P. (2023). Introducing the Dendrify framework for incorporating dendrites to spiking neural networks. *Nature Communications*, 14(1), 131. DOI: https://doi.org/10.1038/s41467-022-35747-8

44. Pan X., Wang Y., Qi Y. (2023). Artificial Neural Network Model and Its Application in Signal Processing. *Asian Journal of Advanced Research and Reports*, 17(1), 1–8. DOI: https://doi.org/10.9734/ajarr/2023/v17i1459

45. Pardede H. F., Adhi P., Zilvan V., Ramdan A., Krisnandi D. (2023). Deep convolutional neural networks-based features for Indonesian large vocabulary speech recognition. *IAES International Journal of Artificial Intelligence*, 12(2), 610. DOI: https://doi.org/10.11591/ijai.v12.i2.pp610-617

46. Pater J. (2019) Generative linguistics and neural networks at 60: Foundation, friction, and fusion. *Language*, 95.1, 41–74.

47. Perrine P. (2023). Inaccessible Neural Language Models Could Reinvigorate Linguistic Nativism. *arXiv preprint arXiv:2301.05272*. DOI: https://doi.org/10.48550/arXiv.2301.05272

48. Recursion. *Cambridge Dictionary*: web site. URL: https://cutt.ly/d7Ob6g7 (date of application: 18.05.23).

49. Reza S., Ferreira M. C., Machado J. J. M., Tavares J. M. R. (2023). A customized residual neural network and bi-directional gated recurrent unit-based automatic speech recognition model. *Expert Systems with Applications*, 215, 119293. DOI: https://doi.org/10.1016/j.eswa.2022.119293

50. Schriner J. (2023). Data Preparation for Fairseq and Machine-Learning using a Neural Network. *Code4Lib Journal*, 55.

51. Selvam K. P., Brorsson M. (2023). DIPPM: a Deep Learning Inference Performance Predictive Model using Graph Neural Networks. *arXiv preprint arXiv:2303.11733*. DOI: https://doi.org/10.48550/arXiv.2303.11733

52. Sethia D., Singh P., Mohapatra B. (2023). Gesture Recognition for American Sign Language Using Pytorch and Convolutional Neural Network. In *Intelligent Systems and Applications: Select Proceedings of ICISA 2022* (pp. 307–317). Singapore: Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-19-6581-4_24

53. Sharma S., Saxena V. P. (2023). Hybrid Sign Language Learning Approach Using Multi-scale Hierarchical Deep Convolutional Neural Network (MDCnn). In *Sentiment Analysis and Deep Learning: Proceedings of ICSADL 2022* (pp. 663–677). Singapore: Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-19-5443-6_51

54. Sonderegger M. (2023). Regression modeling for linguistic data. Cambridge, MA: MIT Press.

55. Suhartono D., Purwandari K., Jeremy N. H., Philip S., Arisaputra P., Parmonangan I. H. (2023). Deep neural networks and weighted word embeddings for sentiment analysis of drug product reviews. *Procedia Computer Science*, 216, 664–671. DOI: https://doi.org/10.1016/j.procs.2022.12.182

56. Teodorescu M. H. (2017). Machine Learning Methods for Strategy Research. *HBS Working Paper* 18-011. Harvard Business School, 59 p. Available at: https://cutt.ly/W6t8UgZ (accessed 18 May 2023).

57. Tereshchenko L., Hlado S. (2022) Nepravdyvist pysmovoho tekstu: shliakhy yii vyznachennia [Falsity of a written text: ways to determine it]. *Psycholinguistics*, 31 (2), 116–136. Available at: https://cutt.ly/Z5FFugk (accessed 18 May 2023). (in Ukrainian)

58. Tuffery S. (2023). Deep Learning for Natural Language Processing. In *Deep Learning: From Big Data to Artificial Intelligence with R* (pp. 431–478). DOI: https://doi:10.1002/9781119845041.ch9

59. Umer M., Imtiaz Z., Ahmad M., Nappi M., Medaglia C., Choi G. S., Mehmood A. (2023). Impact of convolutional neural network and FastText embedding on text classification. *Multimedia Tools and Applications*, 82(4), 5569–5585. DOI: https://doi.org/10.1007/s11042-022-13459-x

60. Veličković P. (2023). Everything is connected: Graph neural networks. *Current Opinion in Structural Biology*, 79, 102538. DOI: https://doi.org/10.1016/j.sbi.2023.102538

61. Wang J., Xie H., Wang F. L., Lee L. K., Wei M. (2023). Jointly modeling intra-and inter-session dependencies with graph neural networks for session-based recommendations. *Information Processing & Management*, 60(2), 103209. DOI: https://doi.org/10.1016/j.ipm.2022.103209

62. Wu J., Yang B., Li D., Deng L. (2023). A semantic relation-aware deep neural network model for end-to-end conversational recommendation. *Applied Soft Computing*, 132, 109873. DOI: https://doi.org/10.1016/j.asoc.2022.109873

63. Wu L., Chen Y., Shen K., Guo X., Gao H., Li S., Long B. (2023). Graph neural networks for natural language processing: A survey. *Foundations and Trends® in Machine Learning*, 16(2), 119–328. DOI: https://doi.org/10.1561/2200000096

64. Zhang G., Wang W., Zhang C., Zhao P., Zhang M. (2023). HUTNet: An Efficient Convolutional Neural Network for Handwritten Uchen Tibetan Character Recognition. *Big Data*. DOI: https://doi.org/10.1089/big.2021.0333