

## PHILOLOGY AND LINGUISTICS

### Specific Features and Distinctive Aspects of Machine Learning in Modern Computational Linguistics AMID SHIVA- and WARMI-world

**Svitlana Goncharenko<sup>1</sup>**

<sup>1</sup>Senior Lecturer, Department of philology and translation;  
Kyiv National University of Technologies and Design, Ukraine

**Abstract.** The article considers the features of the application of machine learning technologies in modern machine linguistics in the conditions of uncertainty, turbulence and systemic risks (SHIVA and WARMI concepts). The main directions of using machine learning methods in natural language processing, their advantages, limitations and development prospects are analyzed. The role of intelligent algorithms in the digital transformation of modern philology is shown.

**Keywords:** modern philology, machine linguistics, natural language processing, machine learning, artificial intelligence, SHIVA-world, WARMI-world.

---

#### Introduction.

Modern philological science is experiencing a period of deep digital transformation [1], caused by the rapid development of Natural Language Processing, Data Mining, Big Data and other innovative technologies of computational linguistics. In the conditions of a constant increase in the volume of digital text corpora [2], the diversity of language data and the complication of communicative processes, the importance of machine learning technologies that allow automating the analysis, interpretation, modeling and generation of natural language is significantly increasing. In such difficult conditions, machine learning is becoming one of the key technological foundations of modern science and practice, ensuring the transition from rigidly formalized algorithms to self-learning intelligent systems [3]. The relevance of the research is further enhanced by significant changes in the environment of the functioning of science and education. Modern society is characterized by a high degree of uncertainty, nonlinearity, turbulence and a multiplicity

## PHILOLOGY AND LINGUISTICS

of interrelated risks [4]. If earlier the concept of VUCA-world was widely used to describe such conditions, today the concepts of SHIVA-world and WARMI-world are becoming increasingly widespread, reflecting a qualitatively new level of complexity of socio-economic, technological and information processes. In such conditions, the need for the use of intelligent methods of information analysis is growing, capable of adapting to rapidly changing data, taking into account uncertainty, revealing hidden patterns and ensuring the stability of the functioning of digital systems [5].

The peculiarity of modern machine linguistics is that language is one of the most complex self-organizing systems, possessing a high degree of ambiguity, variability, context dependence and dynamic variability [6]. That is why traditional algorithmic methods are gradually giving way to machine learning methods, which are able to automatically obtain patterns directly from data without the need for explicit programming of each rule.

In recent years, machine learning methods have significantly changed almost all areas of machine linguistics. They are actively used in automatic translation, text tone analysis, speech recognition, information search, automatic document summarization, knowledge acquisition, text generation, creation of intelligent dialog systems, analysis of large language corpora, and many other tasks of modern digital philology.

That is why the purpose of this article is to study the features, specifics, and current trends in the application of machine learning technologies in machine linguistics, taking into account the concepts of the SHIVA and WARMI worlds, as well as to determine the prospects for the further development of intelligent language technologies in the conditions of high uncertainty of the modern digital environment.

### **Main Part.**

**Machine learning as a technological basis of modern machine linguistics.** Machine learning is a branch of artificial intelligence, based on the construction of mathematical models capable of automatically detecting dependencies between data objects and improving the quality of decisions made as experience accumulates. Unlike classical programming, where the rules are set in advance by the developer, machine learning algorithms independently form internal models based on training samples. For modern machine

## PHILOLOGY AND LINGUISTICS

linguistics, this approach turned out to be particularly effective, since it is almost impossible to completely formalize natural language using a fixed set of grammatical and semantic rules. Real language constructions are constantly changing, contain numerous exceptions, ambiguities and context dependencies. Machine learning allows you to take into account such complexity much more efficiently than traditional algorithms. Currently, machine learning is becoming the central computational paradigm of intelligent language information processing and is the basis of most modern NLP systems.

### ***Main types of machine learning in linguistic research.***

In modern machine linguistics, almost all the main directions of machine learning are used. Supervised learning is applied in the presence of marked-up text corpora and is used for document classification, analysis of emotional coloring of messages, determination of thematic affiliation of texts, automatic recognition of named entities, machine translation and syntactic analysis. Unsupervised learning is aimed at identifying hidden structures within large arrays of text information. It is widely used for thematic modeling, document clustering, identification of semantic groups of words, study of language changes and construction of distributed representations of vocabulary. Semi-supervised learning combines the use of a limited amount of marked-up data with a large volume of unmarked corpora, which is especially important for languages with limited digital resources. Reinforcement learning is gradually beginning to be used in intelligent dialogue systems, interactive machine translation, automatic construction of dialogue strategies and adaptive text generation.

### ***Features of the application of machine learning in natural language processing.***

The main feature of linguistic data is their high level of uncertainty. The same word can have several meanings, the content of a sentence is determined by a wide context, and the interpretation of the text depends on cultural, social, historical and pragmatic factors. Unlike numerical data, text information has a complex hierarchical structure. Morphology, syntax, semantics, pragmatics and discourse are interconnected, forming a multidimensional linguistic space. Therefore, modern machine learning algorithms must take into account not only statistical patterns, but also deep semantic relationships between language units. Additional complexity

## PHILOLOGY AND LINGUISTICS

is created by the constant emergence of new terms, changing language norms, the development of Internet communication, the spread of multimodal messages and the mixing of different language styles. All this requires high adaptability of the models being trained.

### ***Modern machine learning algorithms in machine linguistics.***

The development of computing technologies has led to the widespread use of various machine learning algorithms. Classical statistical models are gradually being supplemented by decision trees, random forests, support vector methods, gradient boosting, probabilistic models, and ensemble algorithms. Deep learning methods have been the most developed. Artificial neural networks allow modeling complex nonlinear dependencies between language units and significantly improve the quality of solving most problems [7]. Transformer architectures occupy a special place, which provide effective modeling of long contexts, allow taking into account complex semantic dependencies between words, and are the basis of modern large language models. Due to this, the quality of automatic translation, text generation, information search, and intelligent document analysis has significantly improved.

### ***Specificity of machine learning in the conditions of SHIVA and WARMI worlds.***

The concepts of SHIVA- and WARMI-world reflect the modern environment, characterized by high dynamics of changes, systemic uncertainty, information overload, technological instability and a multitude of interconnected risks. In such conditions, the requirements for intelligent language systems increase significantly. They must have the ability to constantly adapt, quickly retrain, process streaming data, detect new language patterns and operate stably even with significant changes in input information. Machine learning provides such adaptability due to the ability to continuously update models as new data arrives. This is especially important when analyzing modern information flows, social networks, digital media, scientific publications and news resources, where the language situation changes almost daily.

### ***Advantages of machine learning in modern machine linguistics.***

The use of machine learning technologies provides a number of significant advantages. First of all, the quality of automatic text analysis is significantly improved due to the ability of algorithms to detect complex statistical

## PHILOLOGY AND LINGUISTICS

dependencies that are inaccessible to traditional expert systems. High scalability allows for the effective analysis of corpora with millions of documents. Self-learning of models ensures a constant improvement in the quality of work when accumulating new data. The flexibility of algorithms allows for the use of unified mathematical approaches to solving various linguistic problems. An important advantage is the ability to integrate machine learning with big data technologies, intelligent analytics, generative artificial intelligence, knowledge graphs, and modern digital platforms.

### ***Limitations and current problems in the application of machine learning.***

Despite the obvious advantages, machine learning technologies have a number of significant limitations. First, the effectiveness of most models directly depends on the volume and quality of training data. Second, complex neural networks have limited interpretability, which complicates the explanation of decisions made. Third, there is the problem of transferring models between different languages, especially languages with limited digital corpora. Additional challenges are the computational complexity of training large models, the presence of algorithmic shifts, ethical aspects of using artificial intelligence, copyright protection and ensuring information security of language data. In the conditions of the SHIVA-world, the stability of models to incomplete data, the ability to counteract disinformation and ensuring the reliability of intelligent language systems are of particular importance.

### ***Prospects for the further development of machine learning in digital philology.***

The further development of machine linguistics will be associated with a deeper integration of machine learning, generative artificial intelligence, multimodal models, cognitive computing and intelligent systems for supporting scientific research. Active development of self-learning language models, continuous learning without complete overtraining, federated machine learning, understandable artificial intelligence, hybrid neurosymbol architectures and specialized models for the humanities is expected. Of particular importance will be the creation of adaptive intelligent linguistic platforms capable of functioning in the conditions of constant variability of the information environment, high uncertainty, and systemic risks of modern digital civilization.

## PHILOLOGY AND LINGUISTICS

### Conclusions.

The analysis shows that machine learning has become one of the key technological foundations of modern machine linguistics and plays a decisive role in the digital transformation of modern philology. Its use ensures the transition from traditional algorithmic methods of language processing to intelligent self-learning systems capable of effectively analyzing complex, heterogeneous, and constantly changing language data.

It has been established that the specifics of the application of machine learning in machine linguistics are determined by the high degree of uncertainty of natural language, the multi-level organization of language structures, the contextual dependence of meanings, and the need to process large arrays of text information. In these conditions, the greatest efficiency is demonstrated by modern deep learning methods, transformer architectures, and large language models that ensure high quality analysis, interpretation, and generation of texts.

It is shown that in the conditions of SHIVA and WARMI concepts, adaptability, stability, and intellectual flexibility of machine learning models become of particular importance. Promising areas of further research include the development of understandable artificial intelligence, continuous and federated learning, hybrid neurosymbol models, multimodal language systems, and intelligent platforms for digital philology capable of functioning effectively in the conditions of dynamic uncertainty, turbulence, and systemic risks of the modern information environment.

### References:

- [1] Goncharenko S. (2025). Intelligent information technologies for innovative management of advanced philology projects. Innovations and New Directions in Scientific Research : Proceedings of the 2nd International Scientific Conference (Manchester, United Kingdom, 20 September 2025). – Lulu Press, Inc., 2025. – P. 176–179. URL: <https://er.knutd.edu.ua/handle/123456789/31648>
- [2] M. Krasnyuk, T. Motsyuk, S. Krasniuk (2021). Relevanz des wahlfaches «Wissensdetektion in unstrukturierten daten» bei der ausbildung von mastern der technischen und humanitären fachrichtungen. Education and science of today: intersectoral issues and development of sciences : collection of scientific papers "ΛΟΓΟΣ" with proceedings of the I International scientific and practical conference, Cambridge, United Kingdom, March 19, 2021. Vol. 3. Cambridge-Vinnytsia : P.C. Publishing House & European Scientific Platform, 2021. pp. 66–67.

## PHILOLOGY AND LINGUISTICS

- [3] Krasnyuk, M., & Krasniuk, S. (2021). Modern practice of machine learning in the aviation transport industry. Scientific bulletin АГОРОΣ, May 2021. <https://doi.org/10.36074/logos-30.04.2021.v1.63>
- [4] Nevmerzhytska S. M. (2018). Formation of a strategy for the innovative development of enterprises in conditions of uncertainty. Scientific Bulletin of the Kherson State University. Series: Economic Sciences. 2018. Vol. 32. pp. 99–103. URL: <https://ej.journal.kspu.edu/index.php/ej/article/view/422/418>.
- [5] I. Hrashchenko, S. Krasniuk, T. Tsalko, S. Nevmerzhytska, N. Liubymova (2024). Effective computer modeling in educational management. ScientificWorldJournal. – September 2024. – Issue 27, Part 3. – P. 3–17. – URL: <https://www.sworldjournal.com/index.php/swj/issue/view/swj27-03>
- [6] Maxim Krasnyuk, Svitlana Krasniuk (2024). CHAPTER 6. Evolutionary technologies and genetic algorithms in machine translation. Innovation in modern science: Education and Pedagogy, Philosophy, Philology, Art History and Culture, Medicine and Healthcare. Monographic series «European Science». Book 30. Part 3. 2024. pp. 91–98, Published by: ScientificWorld-NetAkhatAV, Lußstr. 1376227 Karlsruhe, Germany <https://desymp.promonograph.org/index.php/sge/issue/view/sge30-03/sge30-03>
- [7] Maksym Naumenko (2024). Regression analysis using shallow artificial neural networks in the management of an efficient and competitive enterprise. Věda a perspektivy, 7(38) (2024), pp. 17–32. [https://doi.org/10.52058/2695-1592-2024-7\(38\)-17-32](https://doi.org/10.52058/2695-1592-2024-7(38)-17-32)