

PHILOLOGY AND LINGUISTICS

Specifics and Adaptation of Machine Learning in Modern Humanities AMID Turbulence, Uncertainty and Systemic Risks: Evidence from Modern Linguistics

Svitlana Krasniuk¹

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

Abstract. The article explores the specifics, features and directions of adapting machine learning technologies in modern humanities using the example of linguistics. The possibilities of ML for analyzing linguistic data are considered, as well as the problems of uncertainty, turbulence, systemic risks, algorithmic bias and interpretation of results. The need to transition to adaptive machine learning models that combine technological capabilities, expert humanitarian interpretation and a risk-oriented approach is substantiated.

Keywords: *Humanities; Modern linguistics; Computational Linguistics; Machine learning; Artificial intelligence; Natural language processing; Language data; Uncertainty; Turbulence; Systemic risks; Adaptive technologies; Explainable AI; Digital transformation.*

Introduction.

The modern development of the humanities takes place in conditions of rapid complication of the information environment [1], acceleration of digital transformation [2] and constant growth of volumes of text, language and multimodal data [3]. For modern linguistics this means a transition from traditional methods of analysis to the integration of digital, statistical and intellectual technologies capable of processing large arrays of language information [4].

A special place among such technologies is occupied by Machine Learning (ML) technologies, which provide automatic detection of patterns, classification, prediction and modeling of complex processes [5].

However, the application of machine learning in the humanities is fundamentally different from its use in many

PHILOLOGY AND LINGUISTICS

technical and natural science fields [6]. Language is a dynamic, context-dependent, socially and culturally determined system in which the meaning of a linguistic unit depends not only on its formal characteristics, but also on the communicative situation, discourse, cultural context and goals of communication. Therefore, the mechanical transfer of universal ML models to linguistic practice does not provide the necessary level of scientific validity and interpretability [7].

Additional complexity is formed by the modern conditions of the dynamics of uncertainty, turbulence and systemic risks [8]. The variability of the information environment, the emergence of new language practices, the rapid updating of digital content, the spread of multimodal communication, information crises and the instability of data sources lead to the fact that the trained models can quickly lose their relevance [9]. In these conditions, there is a need not just to implement machine learning, but to systematically adapt it to the changing nature of humanitarian knowledge [10].

The relevance of the research is determined by the need to form a scientifically sound approach to the use of machine learning technologies in modern linguistics, which simultaneously takes into account the technological capabilities of ML, the specifics of linguistic data, the requirements of humanitarian interpretation and the impact of uncertainty, turbulence and systemic risks.

Therefore, an important goal of this publication is to identify the specifics and features of the application of machine learning technologies in modern humanities and to substantiate the directions of their adaptation to conditions of dynamic uncertainty, turbulence, and systemic risks using the example of modern linguistics.

Main Part.

1. Specificity of machine learning in modern humanities. Machine learning in the humanities should be considered not as a set of algorithms for automatic data processing, but also as a tool for intellectual support of research activities. Its fundamental difference lies in the need to combine computational efficiency with a meaningful interpretation of the results obtained. In natural science and technical tasks, the quality of a model can often be assessed mainly by quantitative indicators of accuracy, completeness, error or stability of prediction. In the

PHILOLOGY AND LINGUISTICS

humanities, this is not enough. Even a statistically highly accurate model can produce a scientifically limited result if it does not take into account the context, semantics, cultural specificity or communicative function of the material under study. Therefore, the specificity of ML in the humanities is determined by the need to implement the principle of "data - model - interpretation". Machine learning should not replace humanities research, but expand its analytical capabilities, ensuring the detection of hidden patterns and processing of data scales inaccessible to traditional methods.

2. *Features of language data as an object of machine learning in the humanities and modern linguistics.* Linguistic data have a number of characteristics that significantly affect the choice and effectiveness of ML models. These include high variability, ambiguity, contextuality, dependence on discourse, the constant emergence of new language units and changes in the meanings of existing ones. Data from social networks, digital communications, spoken language and multimodal content are especially complex. They are characterized by incompleteness, noise, orthographic variability, language mixing, the use of slang, abbreviations, emojis and other nonlinear means of communication. Therefore, training linguistic models requires not only large volumes of data, but also their high-quality preparation, representativeness, balance and constant updating. The problem of corpus shift - changes in the characteristics of the data on which the model was trained, compared to the data of real operation - also becomes significant.

3. *The main areas of application of ML in the humanities and modern linguistics.* Modern machine learning technologies are used in almost all key areas of computational and applied linguistics. The most significant are automatic text classification, sentiment analysis, topic modeling, named entity recognition, authorship analysis, plagiarism detection, automatic translation, information retrieval, and intelligent processing of large text corpora. A separate area is related to Natural Language Processing (NLP), where ML provides automatic representation and analysis of syntactic, semantic, and pragmatic information. The current stage is characterized by the transition from relatively simple statistical models to neural architectures and large language models. An important perspective is the integration of ML

PHILOLOGY AND LINGUISTICS

with corpus linguistics. This allows the researcher to move from the analysis of limited samples to the study of large-scale dynamic corpora, revealing changes in vocabulary, semantics, discursive practices, and language behavior patterns.

4. *Machine learning under uncertainty in the humanities and modern linguistics.* Uncertainty in modern linguistics is manifested primarily in the unpredictability of language dynamics and incomplete information. The model is trained on a certain historical slice of data, but the language system continues to change. As a result, there is a risk of reducing the quality of predictions when the studied environment changes. Adaptation of ML to uncertainty involves the use of ensemble models, probabilistic methods, regular retraining, estimation of confidence intervals, analysis of error distribution and mechanisms for detecting new data types. Human-in-the-loop – the inclusion of the researcher in the decision-making loop of the model – plays a significant role. Thus, in conditions of uncertainty, not a static, but an adaptive machine learning model becomes optimal, capable of taking into account changes in input data and correcting its own results.

5. *Machine learning under turbulence in the humanities and modern linguistics.* Turbulence is characterized by high speed and unpredictability of changes. In the language environment, it can be caused by social crises, political events, technological changes, information campaigns and the rapid spread of new forms of communication. In such conditions, the traditional scheme of “train a model once and use it for a long time” becomes insufficient. It is necessary to transition to continuous monitoring of the quality of models, dynamic updating of training samples and detection of concept drift. Online learning, incremental learning and adaptive learning are of particular importance. They allow models to be gradually adapted to new data without completely building the entire system from scratch. For linguistics, this is especially important when monitoring new terms, neologisms, discursive changes and transformations of social communication.

6. *Systemic risks of using ML in the humanities and modern linguistics.* The use of machine learning creates not only new opportunities, but also a set of systemic risks. These include the dependence of results on the quality of the source data, algorithmic bias, reproduction of social

PHILOLOGY AND LINGUISTICS

stereotypes, opacity of complex models, loss of contextual information, classification errors and excessive trust of the researcher in automated results. Interpretability is a particular problem. For humanitarian research, it is important not only to obtain a result, but also to understand why the model reached a certain conclusion. Therefore, Explainable AI (XAI) is of particular importance for linguistics. Another risk is the technological dependence of research practice on a limited set of models, platforms and corpora. This can reduce the reproducibility of research and create a dependence of the scientific result on a specific technological infrastructure.

7. *Adaptation and prospects of hybrid machine learning technologies in the humanities and modern linguistics.* A promising direction for the development of machine learning in the humanities is the transition from autonomous algorithmic solutions to hybrid technologies that combine ML computing capabilities, expert knowledge and humanitarian interpretation [11]. Such a model involves active interaction between a person and an intelligent system at the stages of data preparation, training, verification and interpretation of results. This approach is especially significant for modern linguistics and philology, where language data is highly variable, contextual and dynamic. The combination of automated analysis of large corpora with expert close reading allows to simultaneously provide the scale of research and depth of interpretation. The combination of ML, NLP, large language models, Explainable AI and interactive machine learning is also becoming promising. In conditions of uncertainty, turbulence and systemic risks, hybrid systems provide an adaptive loop “data → ML model → expert → correction → new data”, which allows to quickly detect errors, update models and take into account new language phenomena. The human-in-the-loop principle reduces the risk of uncritical trust in automated results and preserves the key role of the researcher. Thus, the prospects of hybrid ML technologies in the humanities are associated not with the complete automation of research, but with the formation of a hybrid intellectual paradigm, in which machine learning provides scale, speed and detection of patterns, and humans – the formulation of research tasks, critical assessment and meaningful interpretation. For philology, this creates opportunities to combine the scale of Big Data with the depth of humanitarian analysis and increase the sustainability of

PHILOLOGY AND LINGUISTICS

research in a dynamic information environment.

Conclusions.

The analysis shows that machine learning technologies are becoming one of the most important tools for the digital transformation of modern humanities, and especially linguistics. Their application allows us to significantly expand the scale and depth of linguistic data research, automate routine analytical operations, identify hidden patterns, and provide processing of large dynamic corpora.

The specificity of ML in the humanitarian sphere is determined by the need to combine high computational efficiency with contextual, semantic, and cultural interpretation of results. In modern linguistics, machine learning should be considered as a replacement for the researcher; its most promising function is to intellectually expand human capabilities.

In conditions of uncertainty, the main requirement is the resistance of models to incompleteness and variability of data; in conditions of turbulence – the ability to quickly update and adapt; in conditions of systemic risks – transparency, controllability, interpretability, and the possibility of expert verification of results.

It is substantiated that the promising adaptation of ML technologies in modern linguistics should be based on the integration of five components: data, algorithms, humanitarian interpretation, risk management and ethical and organizational support. Continue learning, transfer learning, adaptive learning, methods for detecting concept drift and Explainable AI technologies are of particular importance.

In general, the transition from static to adaptive machine learning should be considered as one of the key directions for the further development of digital linguistics. Such a transformation will allow to increase the stability and scientific reliability of linguistic research in the conditions of a constantly changing language, information and socio-cultural environment, while preserving the fundamental role of humanitarian interpretation and expert knowledge.

References:

- [1] Smits, T., & Wevers, M. (2023). A multimodal turn in Digital Humanities: Using contrastive machine learning models to explore,

PHILOLOGY AND LINGUISTICS

- enrich, and analyze digital visual historical collections. Digital Scholarship in the Humanities, 38(3), 1267–1280. <https://doi.org/10.1093/llc/fqad008>
- [2] Rettberg, J. W. (2022). Algorithmic failure as a humanities methodology: Machine learning's mispredictions identify rich cases for qualitative analysis. Big Data & Society, 9(2). <https://doi.org/10.1177/20539517221131290>
- [3] Suissa, O., Elmalech, A., & Zhitomirsky-Geffet, M. (2022). Text analysis using deep neural networks in digital humanities and information science. Journal of the Association for Information Science and Technology, 73(2), 268–287. <https://doi.org/10.1002/asi.24544>
- [4] Goncharenko S. (2026). Innovative machine learning in contemporary philology: Science and practice in the SHIVA world. Trends and Prospects for the Development of Science and Education : Proceedings of the 3rd International Scientific Conference (Oxford, United Kingdom, 3 July 2026). – Bookmundo : United Kingdom, 2026. – P. 223–227.
- [5] Naumenko M., Hrashchenko I. (2024). Suchasnyi shtuchnyi intelekt v antykrizovomu upravlinni konkurentnymy pidpriemstvamy ta kompaniamy [Modern artificial intelligence in anti-crisis management of competitive enterprises and companies]. Grail of Science. No. 42. P. 120–137. DOI: <https://doi.org/10.36074/grail-of-science.02.08.2024.015>
- [6] Krasnyuk, M., & Krasniuk, S. (2021). Modern practice of machine learning in the aviation transport industry. Scientific bulletin АГОФЭ, May 2021. <https://doi.org/10.36074/logos-30.04.2021.v1.63>
- [7] Zubiaga, A. (2024). Natural language processing in the era of large language models. Frontiers in Artificial Intelligence, 6, Article 1343587. <https://doi.org/10.3389/frai.2023.1343587>
- [8] 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>.
- [9] Belém, C. G., Kelly, M., Steyvers, M., Singh, S., & Smyth, P. (2024). Perceptions of linguistic uncertainty by language models and humans. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (pp. 8467–8502). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2024.emnlp-main.483>
- [10] Grieve, J., Bartl, S., Fuoli, M., Grafmiller, J., Huang, W., Jawerbaum, A., Murakami, A., Perlman, M., Roemling, D., & Winter, B. (2024). The sociolinguistic foundations of language modeling. Frontiers in Artificial Intelligence, 7, 1472411. <https://doi.org/10.3389/frai.2024.1472411>
- [11] Hrashchenko, I. S., Krasniuk, M. T., & Krasniuk, S. O. (2019). Hibrydno-stsenarne zastosuvannia intelektualnykh, oriietovanykh na znannia tekhnolohii, yak vazhlyvyi antykrizovyi instrument lohistychnykh kompanii v Ukraini [Hybrid-scenario application of intellectual, knowledge-oriented technologies as an important anti-crisis tool of logistics companies in Ukraine]. Vcheni zapysky Tavriiskoho Natsionalnoho Universytetu imeni VI Vernadskoho. Seria: Ekonomika i upravlinnia-Scientific notes of Tavri National University named after VI Vernadskyi. Series: Economics and management, 30(69), 121–129.