

**International Science Group**

**ISG-KONF.COM**

**IV**

**INTERNATIONAL SCIENTIFIC  
AND PRACTICAL CONFERENCE  
«TECHNOLOGIES, THEORIES AND DEVELOPMENTS:  
MODERN SCIENTIFIC TEACHING»**

**Valencia, Spain**

**September 23-26, 2025**

**ISBN 979-8-89814-218-6**

**DOI 10.46299/ISG.2025.2.4**

# **TECHNOLOGIES, THEORIES AND DEVELOPMENTS: MODERN SCIENTIFIC TEACHING**

Proceedings of the IV International Scientific and Practical Conference

Valencia, Spain  
September 23-26, 2025

**UDC 01.1**

The 4th International scientific and practical conference “Technologies, theories and developments: modern scientific teaching” (September 23-26, 2025) Valencia, Spain. International Science Group. 2025. 183 p.

**ISBN – 979-8-89814-218-6**

**DOI – 10.46299/ISG.2025.2.4**

## EDITORIAL BOARD

<u>Pluzhnik Elena</u>	Professor of the Department of Criminal Law and Criminology Odessa State University of Internal Affairs Candidate of Law, Associate Professor
<u>Liudmyla Polyvana</u>	Department of accounting, Audit and Taxation, State Biotechnological University, Kharkiv, Ukraine
<u>Mushenyk Iryna</u>	Candidate of Economic Sciences, Associate Professor of Mathematical Disciplines, Informatics and Modeling. Podolsk State Agrarian Technical University
<u>Prudka Liudmyla</u>	Odessa State University of Internal Affairs, Associate Professor of Criminology and Psychology Department
<u>Marchenko Dmytro</u>	PhD, Associate Professor, Lecturer, Deputy Dean on Academic Affairs Faculty of Engineering and Energy
<u>Harchenko Roman</u>	Candidate of Technical Sciences, specialty 05.22.20 - operation and repair of vehicles.
<u>Belei Svitlana</u>	Ph.D., Associate Professor, Department of Economics and Security of Enterprise
<u>Lidiya Parashchuk</u>	PhD in specialty 05.17.11 "Technology of refractory non-metallic materials"
<u>Levon Mariia</u>	Candidate of Medical Sciences, Associate Professor, Scientific direction - morphology of the human digestive system
<u>Hubal Halyna Mykolaiivna</u>	Ph.D. in Physical and Mathematical Sciences, Associate Professor

## TABLE OF CONTENTS

CHEMICAL TECHNOLOGIES AND ENGINEERING		
1.	Корчак М.М. АНАЛІЗ ЕНЕРГЕТИЧНИХ ПОКАЗНИКІВ ПОДРІБНЮВАЧА РЕШТОК КУКУРУДЗИ	8
COMPUTER SCIENCE		
2.	Brown J. GOVAI: TRANSPARENT LARGE LANGUAGE MODEL FRAMEWORKS FOR POLICY ANALYSIS AND REGULATORY DECISION SUPPORT	15
3.	Martinez L. EDUAI: A TRUSTWORTHY FRAMEWORK FOR AUTOMATED STUDENT ASSESSMENT AND PERSONALIZED TUTORING	17
4.	Podshyvalova O.. REVIEW OF THE METHODS OF CLASSIFICATION OF YOGA EXERCISES IMPLEMENTED BY MACHINE LEARNING AND NEURAL NETWORKS	19
5.	Suprun A. ARCHITECTURAL FEATURES OF MODERN LARGE LANGUAGE MODELS	23
6.	Єлі М.Я., Байбуз О.Г. МЕТОДИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ РОЗПОДІЛУ РЕСУРСІВ У ХМАРНИХ ОБЧИСЛЕННЯХ: ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ	27
7.	Єлі М.Я., Байбуз О.Г. ОПТИМІЗАЦІЯ РОЗПОДІЛУ РЕСУРСІВ У ХМАРНИХ ОБЧИСЛЕННЯХ: ПІДХОДИ ТА ЕФЕКТИВНІСТЬ АЛГОРИТМІВ	29
8.	Вакульчик С.О., Байбуз О.Г. ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ МЕТОДУ АНАЛІТИЧНОЇ ІЄРАРХІЇ ШЛЯХОМ ІНТЕГРАЦІЇ З НЕЧІТКОЮ ЛОГІКОЮ	31
9.	Вакульчик С.О., Байбуз О.Г. ПРАКТИЧНА ІМПЛЕМЕНТАЦІЯ ГІБРИДНОЇ ФАНР-СИСТЕМИ ТА ЕМПІРИЧНА ВЕРИФІКАЦІЯ ЇЇ ПЕРЕВАГ ДЛЯ БАГАТОКРИТЕРІАЛЬНОГО ПРИЙНЯТТЯ РІШЕНЬ	33

# ARCHITECTURAL FEATURES OF MODERN LARGE LANGUAGE MODELS

**Suprun Anna**

Master in informatics

Kharkiv National University of Radio Electronics

Research supervisor:

Tvoroshenko Iryna

Ph.D., Associate Professor, Department of Informatics

Kharkiv National University of Radio Electronics

Large Language Models (LLMs) are a class of artificial neural networks specifically designed to process and generate natural language based on statistical analysis of large amounts of text data. The foundation of most modern LLMs is the Transformer architecture.

The key element of the Transformer is the attention mechanism, which enables the model to evaluate the importance of each word within the context of a sentence or even an entire document. This makes it possible to effectively handle long-range dependencies and context, which is crucial for generating coherent text.

The attention mechanism is a component of neural networks that allows the model to selectively focus on relevant parts of the input data during processing. In the context of natural language processing, this means that each token in the input sequence considers its relationships with other tokens, forming a more accurate and context-dependent representation. Attention is formalized through weight coefficients that determine the contribution of each token to the output representation.

The training of large language models usually takes place in two stages:

Stage 1. Pre-training – the model is trained on large text collections (books, articles, web pages) with the task of predicting the next word or filling in masked tokens (masked language modeling, as in BERT).

Stage 2. Fine-tuning – the model is further trained on domain-specific data or with the use of instructions to improve its performance on particular tasks.

Some LLMs additionally undergo a stage of Reinforcement Learning with Human Feedback (RLHF). This approach allows the model's responses to better align with human expectations and helps reduce the number of incorrect or harmful outputs [1, 2].

The Transformer architecture consists of encoder and decoder modules, which can be combined depending on the task:

- Encoder-only (BERT) – models composed solely of encoder modules, used for text analysis tasks such as classification, entity extraction, and semantic search;

- Decoder-only (GPT-3/4, LLaMA) – models composed solely of decoder modules. They are mainly used for text generation, as each new token is predicted based on the preceding ones;

– Encoder-decoder (T5, BART) – models that combine both components. They are applied in tasks such as translation, paraphrasing, and summarization, where one text needs to be transformed into another.

The encoder is the part of the model that transforms input text into numerical vectors while preserving its meaning and structure. Each token (a word, a part of a word, or a character) is mapped to an embedding vector, which is enriched with contextual information through multiple layers of attention and neural blocks. In this way, the model captures relationships among all words in a sentence or paragraph, forming a contextualized representation suitable for analysis, classification, or semantic relation extraction. For example, BERT, which relies only on encoder modules, performs exceptionally well on text analysis tasks such as classification, and sentiment detection.

The decoder is the part of the model responsible for text generation. It takes internal representations (from the encoder module or from previously generated tokens in decoder-only models) and predicts the next token sequentially (autoregressive generation). Each new token takes into account all previously generated ones. The multi-head attention mechanism enables the decoder to focus on different aspects of the text simultaneously, improving output quality. Consequently, such models are used for tasks like translation, summarization, dialogue, and creative story generation.

The encoder-decoder Transformer architecture is illustrated in Figure 1 [3].

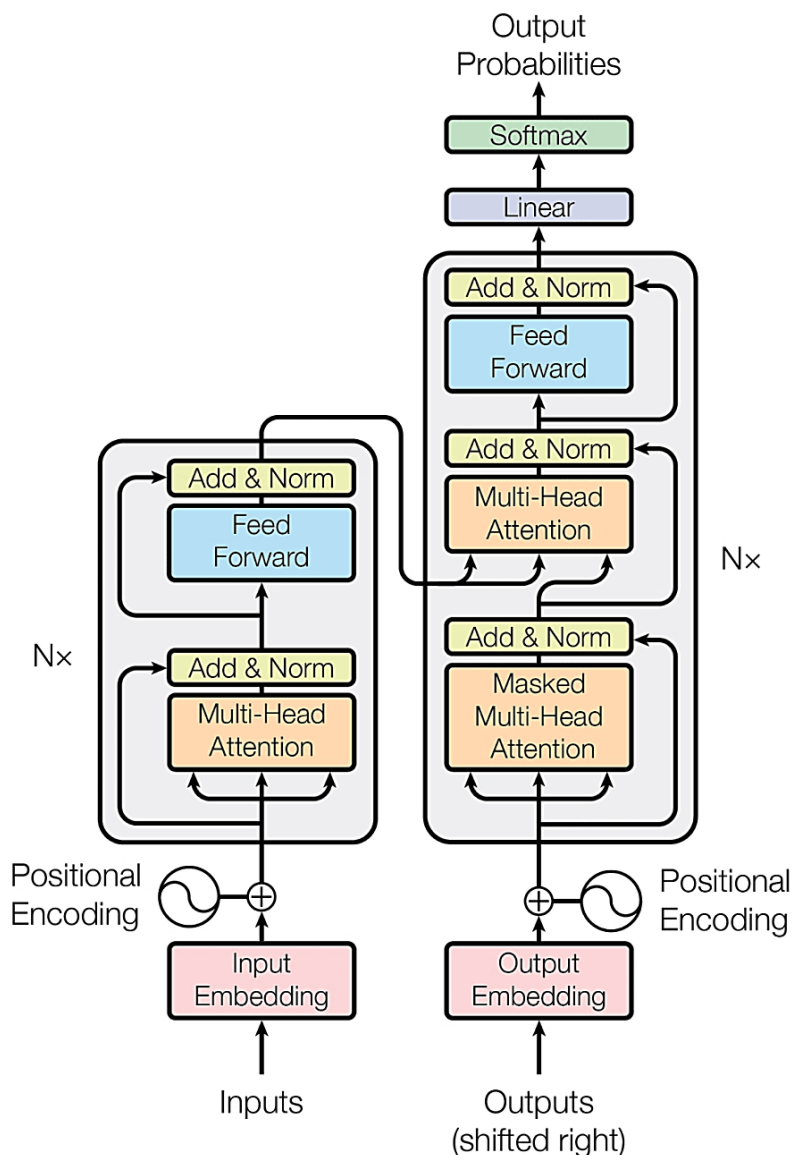
The input text is passed through a sequence of encoder modules, where a contextualized representation of each token is formed. The decoder modules then use these representations to generate the output text by sequentially predicting the next tokens. The figure also shows the self-attention and multi-head attention mechanisms, which enable the model to simultaneously evaluate different aspects of relationships between words, thus ensuring effective handling of long-range dependencies and context.

Large language models can operate in different modes depending on the availability of additional examples or task-specific adaptation:

– Zero-shot learning: the model performs a task without any additional training, relying only on the user's instruction. For example, GPT-4 can immediately translate text from Ukrainian to Japanese without being provided with examples;

– Few-shot learning: the prompt includes a few examples of how the task should be performed, after which the model generates a response in a similar format. This improves accuracy in domain-specific tasks;

– Fine-tuning: further training of the model on a specialized dataset to achieve maximum accuracy in a specific domain (e.g., literary stylization or medical terminology). Fine-tuning can be full (updating all parameters) or partial (e.g., through Low-Rank Adaptation) [4].



**Figure 1** – Transformer architecture (encoder-decoder)

Modern LLMs are evolving toward larger context windows, more efficient attention mechanisms, and multimodality. Models such as Claude 3 or GPT-4 Turbo can process up to 200k tokens of context, which enables analysis of entire books or large codebases. GPT-4o and Gemini 1.5 integrate text, image, audio, and video processing within a single architecture [5-12].

Due to these features, the latest LLMs can effectively work with text both for analysis and for the generation of new plots, characters, and dialogues.

### References:

1. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., Hudáková M., and Gorokhovatskyi O. (2024) Application a committee of Kohonen neural networks to training of image classifier based on description of descriptors set, *IEEE Access*, vol. 12, pp. 73376-73385.
2. Yakovleva O., Matúšová S., Tvoroshenko I., and Isaiev Y. (2024) Visitor counting based on video stream analysis from surveillance cameras to solve various

business problems, *Verejná správa a regionálny rozvoj ekonómia, manažment a marketing*, XX(1), pp. 67-87.

3. Vaswani, A. et al. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

4. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O. (2024) Transforming image descriptions as a set of descriptors to construct classification features, *Indonesian Journal of Electrical Engineering and Computer Science*, 33 (1), 113-125.

5. Гороховатський В., Передрій О., Творошенко І., Марков Т. (2023) Матриця відстаней для множини компонентів структурного опису як інструмент для створення класифікатора зображень, *Сучасні інформаційні системи*, 7(1), С. 5-13.

6. Pomazan, V., Tvoroshenko, I., and Gorokhovatskyi, V. (2023). Development of an application for recognizing emotions using convolutional neural networks, *International Journal of Academic Information Systems Research*, 7(7), pp. 25-36.

7. Gorokhovatskyi, V., Tvoroshenko, I., Kobylin, O., & Vlasenko, N. (2023). Search for visual objects by request in the form of a cluster representation for the structural image description, *Advances in Electrical and Electronic Engineering*, 21(1), pp. 19-27.

8. Daradkeh Y.I., Gorokhovatskyi V., Tvoroshenko I., and Zeghid M. (2024) Improving the effectiveness of image classification structural methods by compressing the description according to the information content criterion, *Computers, Materials & Continua*, vol. 80, no. 2, pp. 3085-3106.

9. Gorokhovatskyi V., Tvoroshenko I. (2023) Identification of visual objects by the search request. *Decision making theory: proceedings of the international symposium, September 28, 2023, Kyiv-Uzhorod, Ukraine*, 25-27.

10. Gorokhovatskyi V., Chmutov Y., Tvoroshenko I., and Kobylin O. (2025) Reducing computational costs by compressing the structural description in image classification methods, *Advanced Information Systems*, vol. 9, no. 1, pp. 5-12.

11. Gorokhovatskyi V., Tvoroshenko I., Yakovleva O., and Hudáková M. (2025) Image description compression in classification structural methods, *IEEE Access*, vol. 13, pp. 43631-43641.

12. Tvoroshenko I., Gorokhovatskyi V., Kobylin O., and Tvoroshenko A. (2023) Application of deep learning methods for recognizing and classifying culinary dishes in images, *International Journal of Academic and Applied Research*, 7(9), pp. 57-70.