МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ

УДК 004

ТЕЗИ ДОПОВІДЕЙ

МІЖНАРОДНОЇ НАУКОВО-ПРАКТИЧНОЇ КОНФЕРЕНЦІЇ "СУЧАСНІ ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ В ЦИФРОВОМУ СУСПІЛЬСТВІ"

16 - 17 квітня 2026

ABSTRACTS OF REPORTS

INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE "MODERN INFORMATION SYSTEMS AND TECHNOLOGIES IN THE DIGITAL SOCIETY"

APRIL 16 - 17, 2026

Харків
2026

Information Systems Department

УДК 004

Матеріали Міжнародної науково-практичної конференції «Сучасні інформаційні системи та технології в цифровому суспільстві»: тези доповідей, 16 – 17 квітня 2026 р. – Х.: ХНЕУ імені Семена Кузнеця, 2026. – 111 с.

Наведені тези пленарних та секційних доповідей за теоретичними та практичними результатами наукових досліджень і розробок. Представлені результати теоретичних та практичних досліджень стосовно галузі комп'ютерних наук, інженерії програмного забезпечення, кібербезпеки, а також систем та технологій інтелектуальної обробки даних.
Матеріали публікуються в авторській редакції.

Materials of the International scientific-practical conference "Modern information systems and technologies in the digital society": abstracts of reports, April 16 - 17, 2026. - Kh.: Simon Kuznets Kharkiv National University of Economics, 2026. – 111 p.

The theses of plenary and sectional reports present theoretical and practical results of scientific research and development. The collection includes findings in the fields of computer science, software engineering, cybersecurity, and intelligent data processing systems and technologies.
The materials are published in the original author's edition.

Disclaimer
The content of these proceedings represents the views of the author only and is his/her sole responsibility. The European Commission does not accept any responsibility for use that may be made of the information it contains

За достовірність викладених фактів, цитат та інших відомостей відповідальність несе автор.

UDC 004.8

Gorokhovatskyi Oleksii
oleksii.gorokhovatskyi@hneu.net

Simyon Kuznets Kharkiv National University of Economics, Kharkiv

SMALL LANGUAGE MODELS FOR PERPLEXITY-BASED TEXT CLASSIFICATION

Classifying whether text is AI-generated or human-written remains a significant challenge, particularly given the rapid development of new large language models (LLMs). Even if a theoretically perfect classifier were available, a further issue arises: what are the explicit reasons underlying its specific decisions?

The use of LLMs, the most powerful artificial intelligence tools, is very popular nowadays. But the majority of them are so large that deployment in local environments, such as a home PC or laptop, is not possible. In this research, we are interested in "small" language models (SLMs) that can run on regular hardware without high resource requirements. As we will be using perplexities, these models should also provide sufficient data to calculate them.

Perplexity is one of the well-known features of the text that reflects the average uncertainty or "surprisal" the model experiences when generating or evaluating text. A low perplexity value indicates that the model confidently chooses the required tokens at generation, a high value indicates a greater variety of tokens used and a higher level of "uncertainty". Conversely, a higher value indicates less reliable predictions of subsequent tokens and poorer model quality. It is believed that human-generated texts have higher perplexity values than AI-generated texts. Language models are optimized to generate text that is statistically probable for their training data, which naturally results in low perplexity by design. Human writing often deviates from these statistical norms, introducing elements of "surprise" or higher perplexity. So, higher perplexity can be interpreted as an indicator of human creativity, beyond statistical features. On the other side, sophisticated LLMs trained on diverse data can reach human-like perplexity.

In this research, we used the dataset proposed and described in [1]. The dataset contained 2533 human-written text chunks, and 2634 AI-generated (GPT-4o-mini) pieces.

At the first stage, we tested different SLMs to determine whether the perplexity calculated on them could be used to classify AI-written and human-written texts. We applied a brute-force search for the classification threshold, and a perplexity value of 14.28 yielded 0.71 accuracy with the Gemma 3 1B IT model; the same accuracy was obtained with the Gemma 3 270m IT model at a perplexity value of 23.28. Finally, the classification based on the Llama 3.2 1B model achieved 0.69 accuracy with a perplexity of 9.28. These models showed positive feedback on the potential of the perplexity value for classification.

Further testing was conducted using a modified version of the dataset, which was built in several steps. The input text is tokenized into numerical IDs and attention masks to distinguish actual tokens from padding; these are then fed into an SLM running in inference mode to produce logits, which are raw prediction scores for each possible next token. The logits are shifted to align with the appropriate input tokens, reflecting the model's next-token prediction mechanism. Finally, a softmax function is applied to the aligned logits, generating output probabilities that provide a detailed, token-level profile of the input text for further analysis. As a result, each input text chunk is represented as a probability vector of size 567 (for Gemma models) and 650 (for Llama). Smaller vectors were padded with zeros to match the length of the longer vector.

Applying KNN to these vectors' classification yielded an accuracy of about 0.68. We have also trained many CNN models with different architectures for classification and selected the best ones. They start with three 1D convolutional layers with 16 RELU-activated neurons and a kernel size of 3, followed by a 1D global average pooling layer. Models based on Gemma 3 SLM had just 2 dense neurons with RELU and an L2 (0.02) regularizer after the feature detection layers and achieved 0.822 and 0.8424 in accuracy for Gemma 3 270m/1 B, respectively. The Llama model features ended with two dense layers containing 16 RELU neurons and 2 RELU neurons. An L2 (0.01) regularizer was applied for both. The accuracy for this model was 0.8559.

It is worth noting that we used small language models (with 1B parameters or fewer) in this research to build feature vectors and classify AI-generated text in Ukrainian. But the text paragraphs we tried to process were generated by much more powerful existing LLM models with hundreds of billions of parameters (gpt-4o-mini for the train/test dataset, GPT5 / Gemini 2.5 Flash / Claude Sonnet 4.5 for the test document example). Probably, implementing the proposed idea could yield better results for the perplexity calculated with more powerful models.

References

1. O. Peredrii. Shallow ANN models to classify Ukrainian AI-generated text. Control, Navigation and Communication Systems, No. 4(82), 2025, pp. 108–113. doi:10.26906/SUNZ.2025.4.108-113.

Information Systems Department

«Сучасні інформаційні системи та технології в цифровому суспільстві»

ТЕЗИ ДОПОВІДЕЙ

Міжнародної науково-практичної конференції “Сучасні інформаційні системи та технології в цифровому суспільстві” 16 – 17 квітня 2026 р.

ABSTRACTS OF REPORTS

International scientific and practical conference "Modern information systems and technologies in the digital society" April 16 - 17, 2026

Відповідальний за випуск: Д.О. Бондаренко

Комп`ютерна верстка: Д.Ю. Голубничий
