

## РЕФЕРАТ

**Актуальність теми.** У сучасному науковому та інформаційному просторі, швидке зростання використання технологій ШІ значно впливає на процеси створення текстового контенту. ШІ може генерувати статті, звіти, аналітичні матеріали, що в зовнішніх ознаках мало чим відрізняються від текстів, написаних людиною. Однак, хоча такі текстові генерації можуть виявитися корисними для зменшення рутинної роботи, вони також породжують нові виклики у контексті автентичності і довіри до наукових та інформаційних джерел. Розвиток технологій, які можуть автоматично ідентифікувати та класифікувати текст, створений за допомогою ШІ, є важливим для забезпечення цілісності наукової комунікації та захисту інтелектуальної власності.

**Об'єктом дослідження** є система організації нейромережевого розпізнавання текстової інформації, згенерованої штучним інтелектом у наукових статтях.

**Предметом дослідження** є методи та алгоритми нейронних мереж, які дозволяють класифікувати текст за допомогою розпізнавання та класифікації його походження.

**Мета роботи:** покращення розпізнавання текстової інформації, згенерованої штучним інтелектом, серед наукових текстів, за допомогою нейронної мережі.

### **Основні завдання включають:**

1. Збір та аналіз даних, що містять як текст, створений людиною, так і ШІ.
2. Розробка і валідація моделі нейронної мережі для класифікації текстів. Навчання та тестування моделі на підготованому наборі даних.
3. Оцінка точності розпізнавання та ефективності моделі.

**Наукова новизна** полягає у створенні новітнього методу для обробки та аналізу наукових текстів, що дозволяє вдосконалити класифікацію та підвищити точність визначення походження тексту.

**Практична цінність** роботи виражається у можливості застосування розробленого методу в академічних та видавничих сферах для верифікації наукового контенту.

**Апробація роботи** основні положення та результати роботи були представлені та обговорювались IV Міжнародна науково-практична конференція “PERSPECTIVES OF CONTEMPORARY SCIENCE: THEORY AND PRACTICE”, 26-28.05.2024, Львів, Україна.

**Структура та обсяг роботи.** Дисертація містить вступ, чотири розділи, висновки і додатки. Описано історію розвитку та теоретичні аспекти використання штучного інтелекту для текстового аналізу, детальний аналіз підходів до навчання мереж, експериментальні методики та оцінку результатів.

*У вступі* надано загальний огляд проблематики дослідження, зосереджено увагу на актуальності розпізнавання тексту, згенерованого штучним інтелектом, в сучасному науковому і практичному контексті. Визначено ключові виклики та висвітлено практичне значення дослідження.

*У першому* розділі проведено детальний аналіз теоретичних аспектів розпізнавання тексту. Розглянуто основні поняття та історію розвитку штучного інтелекту в контексті генерації тексту. Описано сучасні методи та підходи, що використовуються для розпізнавання текстів, зокрема, в області обробки природної мови. Також обговорено переваги та недоліки цих методів.

*У другому* розділі зосереджено увагу на визначенні критеріїв та метрик, які використовуються для оцінювання текстів, згенерованих штучним інтелектом. Детально проаналізовано різні метрики точності та

проведено аналіз типових помилок, які часто зустрічаються під час автоматизованої генерації тексту. Також розглянуто ефективність різних технік на підставі практичних даних, що дозволяє оцінити їхню застосовність у реальних умовах.

У *третьому* розділі проведено аналіз сучасних алгоритмів та методів, які використовуються для розпізнавання тексту, згенерованого ШІ. Обговорено як традиційні методи, так і інноваційні підходи в машинному навчанні та глибокому навчанні. Описано основні алгоритми та їхню адаптацію для конкретних завдань розпізнавання тексту.

У *четвертому* розділі містяться результати реалізації та тестування нейронної мережі, розробленої для розпізнавання ШІ-генерованого тексту. Презентовано методiku валідації, обговорено результати тестувань та порівняно ефективність запропонованого рішення з існуючими методами.

У *висновках* підсумовано основні результати дослідження, підкреслено значення роботи у контексті розвитку технологій розпізнавання тексту та надано рекомендації щодо подальших напрямків розвитку в даній області.

Робота складається з 93 аркушів, включає посилання на літературу.

**Ключові слова:** текстовий аналіз, штучний інтелект, класифікація текстів, нейронні мережі, аутентичність наукових текстів.

## ABSTRACT

**Relevance of the topic.** In the modern scientific and information space, the rapid growth of AI technologies has a significant impact on the processes of creating textual content. AI can generate articles, reports, and analytical materials that are not much different from texts written by humans in their external features. However, while such text generations can be useful for reducing routine work, they also raise new challenges in the context of authenticity and trust in scientific and information sources. The development of technologies that can automatically identify and classify AI-generated text is important for ensuring the integrity of scientific communication and protecting intellectual property.

**The object of research** is a system for organizing neural network recognition of textual information generated by artificial intelligence in scientific articles.

**The subject of the study** is methods and algorithms of neural networks that allow to classify text by recognizing and classifying its origin.

**Aim:** to improve the recognition of textual information generated by artificial intelligence among scientific texts using a neural network.

**The main tasks include:**

1. Collecting and analyzing data containing both human- and AI-generated text.
2. Development and validation of a neural network model for text classification.
3. Training and testing the model on the prepared data set.
4. Evaluation of recognition accuracy and model performance.

**The scientific novelty** is the creation of the latest methods for processing and analyzing scientific texts, which allows to improve the classification and increase the accuracy of determining the origin of the text.

**The practical value of the work** is expressed in the possibility of applying the developed method in the academic and publishing fields to verify scientific content.

### **Testing of the work about**

**Structure and scope of the work.** The thesis consists of an introduction, five chapters, conclusions, and appendices. It describes the history of development and theoretical aspects of using artificial intelligence for text analysis, a detailed analysis of approaches to training networks, experimental methods, and evaluation of results.

*The introduction* provides a general overview of the research topic, focusing on the relevance of AI-generated text recognition in the modern scientific and practical context. The key challenges are identified and the practical significance of the study is highlighted.

*The first* section provides a detailed analysis of the theoretical aspects of text recognition. The basic concepts and history of artificial intelligence development in the context of text generation are considered. Modern methods and approaches used for text recognition, in particular, in the field of natural language processing, are described. The advantages and disadvantages of these methods are also discussed.

*The second* section focuses on the definition of criteria and metrics used to evaluate texts generated by artificial intelligence. Various accuracy metrics are analyzed in detail and typical errors that are often encountered in automated text generation are analyzed. The effectiveness of various techniques is also examined on the basis of practical data, which allows us to assess their applicability in real-world conditions.

*The third* section analyzes modern algorithms and methods used to recognize AI-generated text. Both traditional methods and innovative approaches in machine learning and deep learning are discussed. The main algorithms and their adaptation for specific text recognition tasks are described.

*The fourth* section contains the results of implementation and testing of a neural network developed for recognizing AI-generated text. The validation methodology is presented, the test results are discussed, and the effectiveness of the proposed solution is compared with existing methods.

*The conclusions* summarize the main results of the study, emphasize the importance of the work in the context of the development of text recognition technologies and provide recommendations for further development in this area.

The paper consists of 93 pages and includes references.

**Keywords:** text analysis, artificial intelligence, text classification, neural networks, authenticity of scientific texts