

Міністерство освіти і науки України



NURE

Харківський національний університет
радіоелектроніки

ЗБІРНИК

студентських наукових статей

«Автоматизація та приладобудування»

«Automation and Development of Electronic Devices»

ADED-2025

(Випуск 2)

[електронне видання]



<http://nure.ua/department/kafedra-komp-yuterno-integrovanih-tehnologiy-avtomatizatsiyi-ta-mehatroniki-kitam>



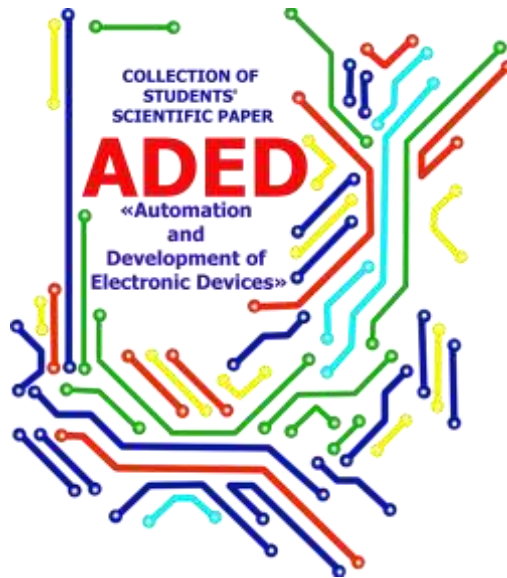
<http://itez.zntu.edu.ua/>



<http://kafea.kdu.edu.ua>

Харків 2025

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки
кафедра комп'ютерно-інтегрованих технологій, автоматизації та робототехніки
(КІТАР)



ЗБІРНИК

студентських наукових статей

«Автоматизація та приладобудування»

«Automation and Development of Electronic Devices»

ADED-2025

(Випуск 2)

[електронне видання]

Харків 2025

ЗМІСТ

<i>Карпович Б.О.</i> Імпульсно-доплерівська селекція в системах автоматичного керування та робототехніці	7
<i>Рожко А.Р., Бондаренко С.В.</i> Підвищення точності систем автоматичного регулювання шляхом корекції динаміки спостерігача стану	12
<i>Бондаренко С.В., Рожко А.Р.</i> Аналіз методів синтезу оптимальних регуляторів для систем із параметричними збуреннями	17
<i>Кобець Д.С., Кравченко С.О.</i> Синтез адаптивних систем із прогнозуючим законом керування	21
<i>Кравченко С.О., Кобець Д.С.</i> Застосування принципу інваріантності для компенсації зовнішніх збурень у системах автоматичного регулювання	25
<i>Коваленко О.А., Бондаренко С.В.</i> Вплив нелінійних характеристик виконавчих механізмів на динамічні властивості систем автоматичного регулювання та методи їх компенсації	29
<i>Lisovskyi A.</i> Comparative Analysis of the Vulnerability of Large Language Models to Prompt Injections	34
<i>Шевченко О.</i> Аналіз методів визначення положення безпілотного наземного мобільного робота на карті місцевості	41
<i>Андреев А. С.</i> Особливості використання LLM в аналізі даних	46
<i>Гайдук І.М.</i> Система управління роботизованим маніпулятором на основі розпізнавання жестів руки	53
<i>Єчевський А. Д.</i> Дослідження ефективності систем навігації SLAM, VSLAM та LDS для автономних мобільних роботів у складських приміщеннях	56
<i>Колбаса О. Р.</i> CRM-система як інструмент інтеграції відділу продажів та виробництва: від зменшення циклу замовлення до підвищення лояльності клієнтів	63
<i>Конєва А. І.</i> Особливості обробки зображень на виробництві	69
<i>Котенко В.А.</i> Аналіз технологій та перспектив розвитку гібридних мобільних роботів	76
<i>Кривчун Р.В.</i> Комп'ютерне моделювання та його роль у сучасному роботизованому виробництві	81
<i>Левченко К.О.</i> Методи кольорового сортування за допомогою контурного виділення звичайною оптичною камерою у видимому спектрі сировини на конвеєрних виробництвах	87
<i>Мамін В.А.</i> Інтелектуальні системи керування квадрокоптерами: аналіз функціональних аспектів та перспективи розвитку	92
<i>Маруніч Р.В.</i>	95

ОСОБЛИВОСТІ ВИКОРИСТАННЯ LLM В АНАЛІЗІ ДАНИХ**Андрєєв А. С.**

Харківський національний університет радіоелектроніки

Україна, 61166, Харків, пр. Науки 14

E-mail: anton.andreiev@nure.ua

Аннотація. Дослідження висвітлює трансформаційний вплив великих мовних моделей (LLM) на процеси аналізу даних. Фокус зроблено на інтеграції LLM у різні етапи data science пайплайну, включаючи підготовку даних, генерацію аналітичних запитів, інтерпретацію результатів та автоматизацію звітності. Робота аналізує переваги впровадження LLM, включаючи прискорення аналітичних процесів, зниження бар'єру входу для нетехнічних користувачів та можливість обробки неструктурованих даних. Особлива увага приділяється можливостям природномовної взаємодії з даними як передовому напрямку розвитку аналітичних систем. Дослідження не обмежується теоретичними аспектами, а надає конкретні приклади застосування LLM у реальних аналітичних сценаріях. Критичний аналіз обмежень та викликів цих технологій забезпечує збалансований погляд на їх потенціал. Візуальні елементи роботи сприяють кращому розумінню складних концепцій. Загалом, дослідження підкреслює ключову роль LLM у демократизації аналізу даних та підвищенні ефективності аналітичних процесів.

Ключові слова: великі мовні моделі, LLM, аналіз даних, обробка природної мови, автоматизація аналітики, data science, штучний інтелект.

FEATURES OF USING LLM IN DATA ANALYSIS**Andreiev A. S.**

Kharkiv National University of Radioelectronics

Ukraine, 61166, Kharkiv, Nauky av.,14

E-mail: anton.andreiev@nure.ua

Abstract. The study highlights the transformative impact of large language models (LLMs) on data analysis processes. The focus is on integrating LLMs into various stages of the data science pipeline, including data preparation, analytical query generation, result interpretation, and reporting automation. The paper analyzes the benefits of implementing LLM, including accelerating analytical processes, lowering the barrier to entry for non-technical users, and the ability to process unstructured data. Particular attention is paid to the possibilities of natural language interaction with data as a cutting-edge direction in the development of analytical systems. The study is not limited to theoretical aspects, but provides specific examples of LLM applications in real-world analytical scenarios. A critical analysis of the limitations and challenges of these technologies provides a balanced view of their potential. Visual elements of the work contribute to a better understanding of complex concepts. Overall, the study highlights the key role of LLM in democratizing data analysis and improving the efficiency of analytical processes.

Key words: large language models, LLM, data analysis, natural language processing, analytics automation, data science, artificial intelligence.

У сучасну цифрову епоху аналіз даних став критично важливим компонентом прийняття рішень у наукових, комерційних та соціальних сферах. Традиційні підходи до аналізу даних вимагають високої технічної кваліфікації та значних затрат часу на написання коду, побудову запитів та інтерпретацію результатів [1-7]. Однак поява великих мовних моделей (Large

Language Models – LLM) відкриває нові горизонти для демократизації аналітичних процесів, дозволяючи ширшому колу користувачів ефективно працювати з даними.

Великі мовні моделі, такі як GPT, Claude, LLaMA та інші, демонструють вражаючі здібності у розумінні природної мови, генерації коду та логічних міркуваннях. Ці можливості роблять LLM потужним інструментом для аналізу даних, дозволяючи користувачам формулювати аналітичні запити природною мовою, автоматично генерувати код для обробки даних, та отримувати зрозумілі пояснення результатів аналізу.

Інтеграція LLM в аналітичні процеси тісно пов'язана з розвитком автоматизації та штучного інтелекту. Автоматизація рутинних завдань з підготовки даних, побудови візуалізацій та генерації звітів може значно підвищити продуктивність аналітиків, дозволяючи їм зосередитися на більш складних та творчих аспектах роботи [8-15]. Використання LLM сприяє формуванню нових, більш інтуїтивних та доступних аналітичних систем, які можуть не тільки відповідати на конкретні запити користувачів, але й пропонувати інсайти на основі контекстуального аналізу даних.

Таким чином, у контексті постійного зростання обсягів даних та зростаючої потреби в швидкому прийнятті рішень, актуальність цього дослідження стає очевидною [16-18]. Інтеграція LLM відкриває нові можливості для покращення швидкості, точності та доступності аналізу даних. У сучасному інформаційному середовищі, де обсяги даних зростають експоненціально, а потреба у швидкому аналізі стає дедалі критичнішою, роль LLM у процесах data science стає ключовою. Великі мовні моделі не лише дозволяють обробляти різноманітні типи даних швидше і з меншими помилками, але й адаптуються до специфічних потреб користувачів, покращуючи якість та доступність аналітичних процесів. В таблиці 1 надано детальний огляд того, як LLM впливають на різні етапи аналізу даних і сприяють їх вдосконаленню.

Таблиця 1 – Роль LLM у процесах аналізу даних

Можливість LLM	Роль у процесах аналізу даних
Розуміння природної мови	Інтерпретація аналітичних запитів, сформульованих природною мовою. Переклад бізнес-питань у технічні завдання аналізу. Зниження бар'єру входу для нетехнічних користувачів.
Генерація коду	Автоматичне створення SQL запитів, Python/R скриптів для аналізу. Прискорення розробки аналітичних пайплайнів. Зменшення ймовірності синтаксичних помилок.
Аналіз та інтерпретація	Виявлення патернів та аномалій у даних. Генерація описових статистик та інсайтів. Пояснення складних аналітичних результатів простою мовою.
Автоматизація звітності	Створення текстових звітів на основі аналітичних результатів. Генерація наративів для візуалізацій. Адаптація звітів до різних аудиторій.

Сучасні системи аналізу даних використовують LLM для вдосконалення широкого спектру аналітичних процесів, від підготовки даних до створення комплексних звітів. LLM демонструють високу ефективність у завданнях підготовки та очищення даних, що традиційно займає до 80 % часу аналітиків. Концепція використання LLM для підготовки даних на рис. 1.

На етапі «Розуміння структури даних» LLM аналізують метадані, зразки записів та документацію для побудови повного розуміння структури датасету. Модель може автоматично визначити типи даних, виявити зв'язки між таблицями та ідентифікувати ключові поля.

На етапі «Виявлення проблем якості даних», використовуючи можливості розуміння контексту, LLM можуть виявляти аномалії, пропущені значення, дублікати та інші проблеми якості даних. Модель може запропонувати стратегії їх вирішення на основі контексту та найкращих практик.

Після ідентифікації проблем LLM генерують відповідний код для очищення даних, використовуючи популярні бібліотеки, такі як pandas, dplyr або SQL, залежно від технологічного стеку користувача.



Рисунок 1 – Концепція використання LLM

Розглянемо приклади застосування LLM у підготовці даних:

- стандартизація форматів – LLM можуть генерувати код для приведення дат, валют та інших форматів до єдиного стандарту, розпізнаючи різні варіанти представлення даних;
- обробка пропущених значень – модель аналізує природу пропусків та пропонує відповідні стратегії імпутації (заповнення середнім, медіаною, forward fill тощо);
- виявлення та обробка викидів – LLM можуть визначити контекстуально доцільні методи виявлення викидів та запропонувати підходи до їх обробки;
- трансформація змінних – автоматична генерація коду для кодування категоріальних змінних, нормалізації числових ознак та створення нових features.

Однією з найпотужніших можливостей LLM є здатність перетворювати природномовні запити на робочий аналітичний код. Цикл обробки запитів за допомогою LLM наведено на рис. 2.



Рисунок 2 – Цикл обробки запитів за допомогою LLM

На етапі семантичного аналізу модель інтерпретує запит користувача, щоб визначити необхідні набори даних, типи їх обробки (фільтрація, агрегація), а також форми подання результату у вигляді графіків чи метрик.

Модель здатна генерувати код, що враховує контекст проекту, такі як існуючі структури даних, змінні та прийнятий в команді стиль кодування. Це забезпечує створення рішень, які органічно інтегруються в поточну робочу платформу. Крім того, моделі можуть аналізувати та покращувати продуктивність запитів шляхом їх оптимізації з урахуванням таких факторів, як наявні індекси, стратегії партиціонування та інші критерії ефективності. Приклади застосування:

- SQL запити з природної мови – користувач може запитати: «Покажи топ-10 клієнтів за виручкою за останній квартал», і LLM згенерує відповідний SQL запит з усіма необхідними JOIN, GROUP BY та ORDER BY конструкціями;

- Python скрипти для статистичного аналізу – за запитом «Проведи A/B тест між контрольною та тестовою групами» LLM створить скрипт з відповідними статистичними тестами, перевірками припущень та візуалізаціями;

- R код для регресійного аналізу – LLM можуть генерувати повні аналітичні пайплайни, включаючи підготовку даних, побудову моделей, валідацію та інтерпретацію результатів;

- запити до NoSQL баз даних – моделі здатні працювати з MongoDB, Elasticsearch та іншими NoSQL системами, генеруючи відповідні запити з урахуванням специфіки їх синтаксису.

LLM вносять значний вклад у найбільш важливий етап аналітичного процесу – інтерпретацію результатів та генерацію практичних рекомендацій. Процес інтерпретації аналітичних результатів за допомогою LLM наведено на рисю. 3.



Рисунок 3 – Процес інтерпретації аналітичних результатів за допомогою LLM

Автоматичне генерування описів – LLM створюють зрозумілі текстові описи числових результатів, що особливо корисно для комунікації з нетехнічними стейкхолдерами. Модель може пояснити, що означають статистичні показники та чому вони важливі.

Виявлення нетривіальних патернів – використовуючи можливості розуміння контексту, LLM можуть виявляти неочевидні зв'язки в даних та звертати увагу аналітиків на потенційно цікаві закономірності для подальшого дослідження.

Контекстуальна інтерпретація – модель враховує бізнес-контекст, історичні дані та доменні знання для надання більш релевантних та практично застосовних інсайтів.

Генерація гіпотез – на основі виявлених патернів LLM можуть пропонувати гіпотези для подальшого тестування та напрямки досліджень. Приклади використання LLM для

інтерпретації:

- пояснення трендів – LLM можуть описати виявлені тренди простою мовою, наприклад: «Спостерігається сезонне зростання продажів на 23 % у четвертому кварталі, що узгоджується з святковим періодом»;

- аналіз кореляцій – модель може пояснити виявлені кореляції: «Сильна позитивна кореляція між витратами на маркетинг та кількістю нових клієнтів ($r = 0,78$) свідчить про ефективність маркетингових кампаній»;

- інтерпретація моделей машинного навчання – LLM можуть пояснювати feature importance, SHAP values та інші метрики інтерпретованості моделей доступною мовою;

- генерація виконавчих резюме – автоматичне створення executive summaries, які виділяють ключові висновки та рекомендації для топ-менеджменту.

Використання великих мовних моделей в аналізі даних приносить суттєві переваги, але також супроводжується певними викликами та обмеженнями, які важливо враховувати при впровадженні цих технологій.

LLM надають унікальні можливості для трансформації аналітичних процесів:

- демократизація аналізу даних – зниження технічного бар'єру дозволяє бізнес-користувачам самостійно проводити базовий аналіз без глибоких знань програмування;

- прискорення аналітичних процесів – автоматизація генерації коду, візуалізацій та звітів може скоротити час виконання аналітичних завдань на 50-70 %;

- підвищення якості аналізу – LLM можуть запропонувати додаткові перевірки, альтернативні підходи та виявити потенційні помилки в логіці аналізу;

- покращена документація – автоматична генерація коментарів до коду, пояснень методології та документації результатів.

Тож, визначимо обмеження LLM в аналізі даних:

1. Галюцинації – LLM можуть генерувати правдоподібний, але неправильний код або інтерпретації, що вимагає постійної верифікації результатів людиною-експертом.

2. Обмеження контекстного вікна – існуючі обмеження на розмір контексту ускладнюють роботу з великими датасетами та складними багатоетапними аналізами.

3. Відсутність глибоких доменних знань – LLM можуть не враховувати специфічні особливості конкретних галузей або бізнес-контексту організації.

4. Складність верифікації – перевірка коректності згенерованого коду та результатів може вимагати стільки ж часу, скільки і написання коду вручну.

5. Проблеми з продуктивністю – згенерований код може бути неоптимальним з точки зору продуктивності, особливо для роботи з великими обсягами даних.

Незважаючи на ці виклики, переваги використання LLM в аналізі даних є суттєвими. Важливо розуміти обмеження технології та застосовувати її розумно, поєднуючи автоматизацію з людською експертизою. LLM найкращим чином працюють як інструменти підтримки аналітиків, а не як їх повна заміна.

Інтеграція великих мовних моделей у процеси аналізу даних відкриває нові можливості для демократизації аналітики та підвищення ефективності роботи з даними. LLM суттєво змінюють ландшафт data science, дозволяючи ширшому колу користувачів взаємодіяти з даними через природну мову та автоматизуючи рутинні аналітичні завдання.

Використання LLM на різних етапах аналітичного процесу – від підготовки даних до генерації інсайтів – демонструє значний потенціал для прискорення аналітичних робіт та покращення якості результатів. Можливості природномовної взаємодії, автоматичної генерації коду та інтелектуальної інтерпретації результатів роблять аналіз даних більш доступним та ефективним. Однак важливо розуміти обмеження цієї технології. Проблеми точності, конфіденційності даних та необхідність верифікації результатів означають, що LLM повинні

використовуватися як інструменти підтримки, а не як повна заміна людської експертизи. Найбільшої ефективності можна досягти при поєднанні потужності LLM з критичним мисленням та доменними знаннями аналітиків.

Таким чином, впровадження LLM в аналітичні процеси не лише оптимізує роботу з даними, але й підвищує загальну культуру data-driven прийняття рішень в організаціях, роблячи аналіз даних більш доступним та ефективним інструментом для всіх користувачів у сучасному інформаційному середовищі. В роботі розкриті різні аспекти використання LLM (підготовка даних, генерація коду, інтерпретація результатів) та надає конкретні приклади їх застосування в реальних аналітичних сценаріях. Критичний аналіз переваг та обмежень забезпечує збалансоване розуміння потенціалу цієї технології.

ЛІТЕРАТУРА:

1. Rudenko, M., et al. Overview of approaches to scaling relational databases in development and adaptation of web applications // Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій: Тези доповідей XII Міжнародної науково-практичної конференції (10-12 грудня 2024 р., м. Запоріжжя). [Електронний ресурс] /Електрон. дані. – Запоріжжя: НУ «Запорізька політехніка», 2024. – С. 398-402
2. Achkan, M. S., et al. Integration of cloud technologies into modern SCADA systems: prospects and challenges // «Computer-integrated technologies, automation and robotics» CITAR-2025. – 2025. – pp. 26-29
3. Konieva, A., et al. Main trends in the development of automated image processing systems // «Computer-integrated technologies, automation and robotics» CITAR-2025. – 2025. – pp. 68-72
4. Rudenko, M., et al. Overview of algorithmic approaches to forecasting in CRM systems // «Computer-integrated technologies, automation and robotics» CITAR-2025. – 2025. – pp. 101-105
5. Sotnik, S., et al. Evaluating relational database scaling strategies in web engineering // International Conference on Advanced Trends In Radioelectronics and Infocommunications (ATRIC-2025) (May 21–22, 2025), Lviv Polytechnic Publishing House, Lviv, Ukraine, 2025. – pp. 224-228
6. Сотник, С. Розробка автоматизованої інформаційно-пошукової системи вибору маніпулятора промислових роботів // Електромеханічні і енергозберігаючі системи. – 2025. – № 1 (68). – С. 52-58
7. Sotnik, S. Development of a range measurement module on an ultrasonic sensor with a GSM module // Radio Electronics, Computer Science, Control. – 2. – 2025. – pp. 32-44
8. Danylenko, M. M., et al. Comparative analysis of modern SCADA packages for production automation // International Journal of Academic Engineering Research (IJAER). – 2025. – Vol. 9. – 2. – pp. 26-34
9. Marunich, R.V., et al. Modern IoT technologies for creating automated access systems // Sustainable smart cities and communities: business and innovation solutions 2025: Proceedings of I st I International Conference, Kharkiv, April 21, 2025: Theses of Reports. – 2025. – pp. 38-39
10. Yechevskiy, A., et al. Analysis of the data collection process about products at different stages of production // Manufacturing & Mechatronic Systems 2025: Proceedings of IX st International Conference, Kharkiv, October 25-26, 2025: Theses of Reports. – 2025. – pp. 38-4
11. Rudenko, M., et al. Classification of CRM systems // Manufacturing & Mechatronic Systems 2025: Proceedings of IX st International Conference, Kharkiv, October 25-26, 2025: Theses of Reports. – 2025. – pp. 42-45
12. Fesenko, A., et al. Review and selection of optimal sensors for building a production facility microclimate monitoring system // Manufacturing & Mechatronic Systems 2025: Proceedings of IX st International Conference, Kharkiv, October 25-26, 2025: Theses of Reports. – 2025. – pp. 50-53

13. Cherednichenko, T., et al. Features of automatic working time control systems // Manufacturing & Mechatronic Systems 2025: Proceedings of IX st International Conference, Kharkiv, October 25-26, 2025: Theses of Reports. – 2025. – pp. 54-57
14. Andreiev, A., et al. “Web application security: protection against modern cyber threats” Overview of key vulnerabilities (XSS, CSRF, SQL injections), protection methods, use of HTTPS, authentication, and authorization // Manufacturing & Mechatronic Systems 2025: Proceedings of IX st International Conference, Kharkiv, October 25-26, 2025: Theses of Reports. – 2025. – pp. 66-70
15. Sotnik, S. Integration of IoT into security systems: opportunities and risks // International Journal of Academic Engineering Research (IJAER), 2024. – Vol. 8, Issue 11. – pp. 56-61
16. Khalimonov, Y. I., et al. Monitoring and optimising conditions in production environment // Proceedings of the XVII International scientific and practical conference «Information technologies and automation– 2024», 2024. – pp. 256-258
17. Lvov, A. A., et al. Overview of digital locks with artificial intelligence // Комп’ютерно-інтегровані технології автоматизації технологічних процесів на транспорті та у виробництві : матеріали всеукр. наук.-практ. конф. здобувачів вищ. освіти і молодих учених, 20 листоп. 2024 р. / Харків. нац. автомоб.-дор. ун-т. – Харків, 2024. – С. 112-116
18. Sotnik, S. V. Development of automated control system and registration of metal in continuous casting // [Radio Electronics, Computer Science, Control](#). – 3. – 2024. – pp. 197-211