

РОЗРОБЛЕННЯ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ АНАЛІЗУ ВХІДНОЇ ІНФОРМАЦІЇ РОБІТНИКА ПРИЛАДОБУДІВНОГО ВИРОБНИЦТВА ДЛЯ ВИДАЧІ ЗАВДАНЬ НА ВИКОНАННЯ

А.С. Андрєєв

Харківський національний університет радіоелектроніки

Україна, 61166, Харків, пр. Науки 14

E-mail: anton.andreiev@nure.ua

Анотація: У статті розглядається розробка програмного забезпечення для аналізу текстової інформації робітників приладобудівного виробництва. Досліджуються методи інтелектуального аналізу тексту, технології обробки природної мови та використання великих мовних моделей для обробки виробничої документації. Представлено архітектуру системи з використанням Python, FastAPI і LLM-моделей, що дає змогу автоматизувати процеси аналізу CV робітників, обробки запитів та генерації персоналізованих завдань на виконання. Описано технічні засоби реалізації та етапи роботи системи від отримання даних до взаємодії з користувачем через веб-інтерфейс.

Ключові слова: приладобудівне виробництво, аналіз текстової інформації, великі мовні моделі, обробка природної мови, Python, FastAPI, машинне навчання, штучний інтелект, класифікація текстів, автоматизація документообігу, LLM, інтелектуальний аналіз тексту.

DEVELOPMENT OF SOFTWARE FOR ANALYZING THE INPUT INFORMATION OF AN INSTRUMENTATION PRODUCTION WORKER TO ISSUE TASKS FOR EXECUTION

A. Andreiev

Kharkiv Kharkiv National University of Radio Electronics Ukraine, 61166, Kharkiv, Nauky av, 14

E-mail: anton.andereiev@nure.ua

Annotation: The article discusses the development of software for analyzing textual information of instrumentation production workers. The article investigates methods of text mining, natural language processing technologies and the use of large language models for processing production documentation. The system architecture using Python, FastAPI, and LLM models is presented, which allows automating the processes of analyzing workers' CVs, processing queries, and generating personalized tasks for execution. The technical means of implementation and the stages of the system's operation from data acquisition to user interaction via the web interface are described.

Key words: instrumentation, text information analysis, big language models, natural language processing, Python, FastAPI, machine learning, artificial intelligence, text classification, document automation, LLM, text mining.

В умовах стрімкого розвитку промисловості та підвищення вимог до якості продукції, впровадження сучасних технологій стає ключовим фактором успіху приладобудівних підприємств[1]. Цифрова трансформація виробничих процесів відкриває нові можливості для оптимізації виробництва, підвищення ефективності та конкурентоспроможності підприємств галузі[2]. Особливу роль у цьому відіграють інтелектуальні системи аналізу даних, які дозволяють обробляти великі масиви вхідної інформації та приймати обґрунтовані управлінські рішення в режимі реального часу.

Сучасні технології, такі як машинне навчання, штучний інтелект та промисловий інтернет речей (ІоТ), стають невід'ємною частиною приладобудівного виробництва[3]. Вони забезпечують автоматизацію процесів збору та обробки даних, підвищують точність

вимірювань, знижують вплив людського фактору та дозволяють своєчасно виявляти потенційні проблеми у виробничому процесі. Це особливо важливо в умовах зростаючої складності приладобудівної продукції та необхідності забезпечення її високої якості та надійності.

В сучасному приладобудуванні аналіз текстової інформації робітників являє собою складний багатоетапний процес, який використовує різноманітні методи та технології обробки даних[4]. На сьогоднішній день можна виділити кілька ключових підходів до аналізу текстової інформації:

1) Методи інтелектуального аналізу тексту (Text Mining):

а) Кластерний аналіз для групування схожих текстових повідомлень та виявлення прихованих закономірностей

б) Класифікація текстових даних для автоматичного сортування звітів та коментарів

в) Аналіз тональності текстів для виявлення проблемних ситуацій

г) Вилучення ключових тем та понять з текстових документів робітників

2) Технології обробки природної мови (NLP):

а) Розпізнавання іменованих сутностей для ідентифікації важливих елементів у текстах

б) Семантичний аналіз для розуміння змісту документації

в) Автоматичне узагальнення великих обсягів текстової інформації

3) Аналіз текстових даних за допомогою великих мовних моделей: Автоматична інспекція якості компонентів

а) Автоматична класифікація текстових звітів

б) Генерація персоналізованих завдань на основі аналізу резюме та запитів

в) Автоматичне визначення пріоритетності та складності завдань відповідно до навичок робітника

Важливим аспектом сучасного приладобудівного виробництва є ефективна обробка текстової інформації, що надходить від робітників виробництва у різних формах: звіти про виконані операції, записи в журналах, пояснювальні записки, текстові коментарі до виконаних завдань, пропозиції щодо вдосконалення процесів[5]. Інтеграція різних методів обробки тексту у єдину систему забезпечує комплексний аналіз всієї текстової інформації, що дозволяє отримати цінні інсайти для оптимізації виробництва.

Вибір технічних засобів для створення програмного забезпечення з аналізу даних робітників базується на використанні найсучасніших технологій штучного інтелекту. Ключовим компонентом системи є великі мовні моделі (LLM), які забезпечують глибокий аналіз текстового контенту у звітах робітників, коментарях до виробничих процесів та пропозиціях щодо покращення роботих[7].

Основою технічного рішення є інтеграція LLM-моделей (таких як GPT або BERT) з системами збору даних на робочих. Це дозволяє:

а) Автоматичний аналіз CV робітників для визначення їх кваліфікації та компетенцій

б) Витягувати ключову інформацію з запитів користувачів

в) Створювати структуровані завдання на виконання з урахуванням кваліфікації робітника

г) Визначати пріоритетність виконання завдань на основі аналізу вхідної інформації

Для забезпечення ефективної роботи системи використовується Python як основна мова програмування та хмарна інфраструктура з інтегрованими GPU-прискорювачами, що дозволяє обробляти великі обсяги текстової інформації. Зберігання та індексація текстових даних реалізується за допомогою спеціалізованих баз даних, оптимізованих для роботи з текстом.

Система забезпечує:

а) Глибокий аналіз CV та запитів користувачів за допомогою LLM

б) Інтелектуальний аналіз кваліфікації робітників

в) Створення персоналізованих завдань на основі вхідної інформації

Важливим аспектом є можливість тонкого налаштування (fine-tuning) LLM-моделей під специфіку текстової документації конкретного приладобудівного підприємства, включаючи[6]:

- а) Адаптацію до технічної термінології підприємства
- б) Навчання на історичних текстових документах та звітах робітників
- в) Налаштування правил класифікації та виділення ключової інформації

Архітектура системи включає проміжний шар безпеки, який забезпечує конфіденційність обробки текстових даних та захист від витоку чутливої інформації про виробничі процеси. Це дозволяє безпечно автоматизувати процеси обробки текстової документації без ризику порушення комерційної таємниці.

Для розробки програмного забезпечення з аналізу текстової інформації робітників на приладобудівному виробництві обрано мову програмування Python через її потужні можливості в області обробки природної мови та інтеграції з системами документообігу. Python має ряд переваг для даного завдання:

- а) Багата екосистема бібліотек для роботи з текстовими даними (pandas, nltk)
- б) Зручні фреймворки для створення веб-інтерфейсу (Flask, FastAPI)
- в) Бібліотеки для інтеграції з LLM-моделями (transformers, langchain) для глибокого аналізу тексту
- г) Широкі можливості для роботи з різними форматами текстових документів (docx, pdf, txt)

Технічні засоби включають:

- а) PostgreSQL для зберігання структурованих даних
- б) Docker для контейнеризації компонентів системи

На етапі отримання даних система може збирати текстову інформацію з різних джерел: електронних документів та текстових файлів. За допомогою FastAPI можна налаштувати інтеграцію з корпоративними системами через API для автоматичного отримання текстових звітів та документації робітників.

Текстові дані потім зберігаються в базі даних, що дає змогу швидко здійснювати пошук і доступ до потрібної інформації. Також, для зручності користувачів системи, є можливість вручну вводити текстові дані або завантажувати текстові файли для обробки. Це дає змогу інтегрувати як автоматизовані, так і ручні методи збору текстової інформації про виробничі процеси.

На етапі опрацювання результатів система аналізує попередньо опрацьовані текстові дані від робітників з використанням LLM-моделей для вилучення глибших патернів, зв'язків і смислової інформації. На цьому етапі текстові дані, які пройшли очищення і нормалізацію, обробляються за допомогою потужних моделей штучного інтелекту, що дають змогу зрозуміти значення і контекст повідомлень робітників навіть у неструктурованих текстах.

LLM-моделі, такі як GPT або BERT, використовуються для розпізнавання складних взаємозв'язків у тексті, виявлення тем і тенденцій, а також для класифікації та категоризації отриманих текстових даних. Наприклад, система може автоматично визначати оптимальні завдання для робітників на основі їхнього CV, враховувати їхню кваліфікацію при розподілі завдань та формувати чіткі інструкції щодо виконання з урахуванням індивідуальних особливостей кожного робітника.

За допомогою таких моделей система може також аналізувати запити користувачів, що дає змогу точніше визначити потреби та очікування, а також формувати завдання з урахуванням специфіки запиту. Це важливо для ефективного розподілу робочого навантаження та підвищення продуктивності робітників.

Результати аналізу можуть бути представлені користувачеві у вигляді зручних звітів через веб-інтерфейс, що дає змогу зручно працювати з результатами. Веб-інтерфейс також надає можливість пошуку за проаналізованими даними, що полегшує прийняття рішень і планування подальших дій на основі отриманих результатів.

На етапі взаємодії з користувачем система надає інтерфейс для зручної роботи з отриманими та обробленими даними. Для цього використовується веб-додаток із фронтендом, розробленим із використанням сучасних технологій який взаємодіє з бекендом на базі FastAPI.

Користувач може завантажувати нові дані (наприклад, файли або текст), які система автоматично обробляє на серверній стороні через FastAPI. Завдяки інтеграції з LLM-моделями, система може надавати більш точні та корисні рекомендації на основі аналізу тексту. Веб-інтерфейс дає змогу відображати звіти, що допомагає користувачеві краще зрозуміти та інтерпретувати результати аналізу.

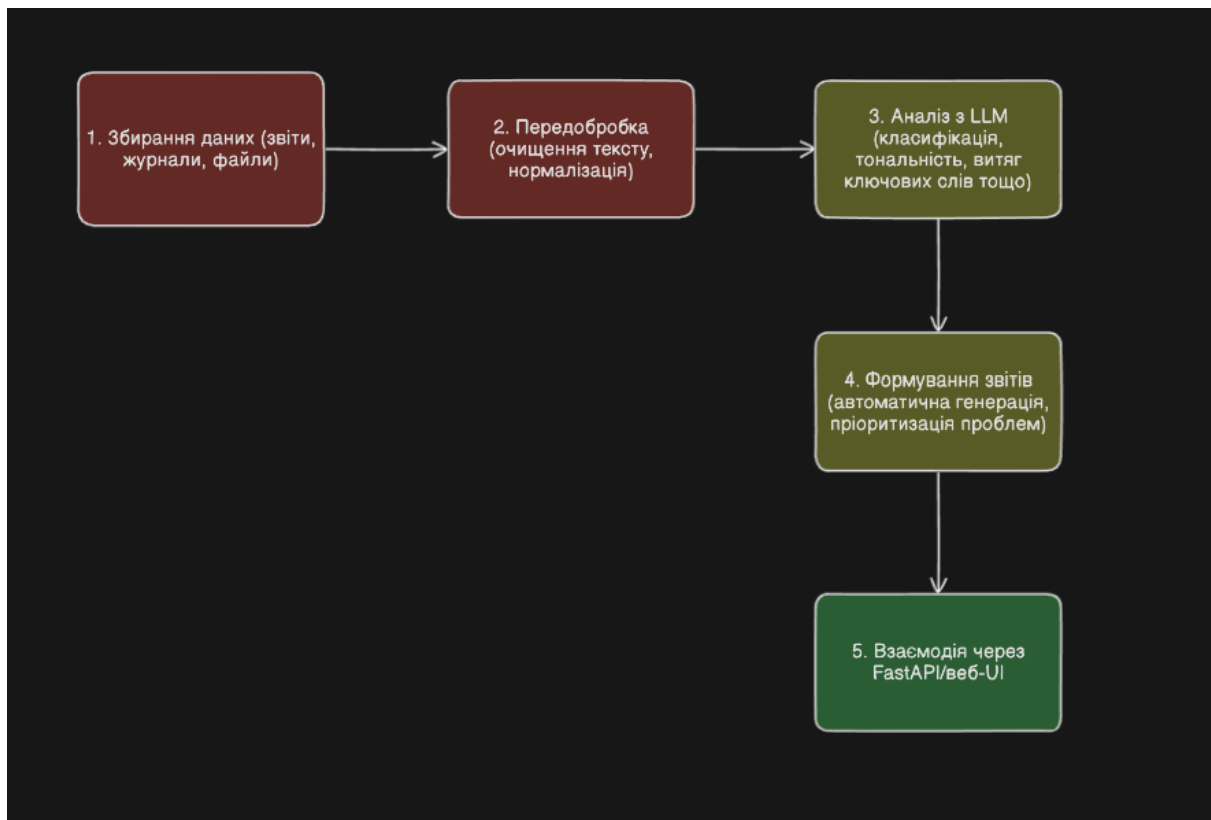


Рисунок 1 – Архітектура системи аналізу текстової інформації

Таким чином, розроблена система на базі Python, FastAPI та LLM-моделей надає потужні інструменти для обробки, аналізу та взаємодії з текстовими даними робітників приладобудівного виробництва. Вона дає змогу ефективно автоматизувати процеси обробки текстової інформації, аналізувати її зміст, виявляти проблеми та тенденції, а також надавати користувачеві зручний та інтуїтивно зрозумілий інтерфейс для перегляду результатів. Такий підхід значно прискорює ухвалення рішень і покращує взаємодію між користувачем і системою, що своєю чергою підвищує ефективність роботи на підприємстві.

ЛІТЕРАТУРА:

1 Johnson, M., & Smith, K. (2023). "Modern Technology Integration in Instrumentation Manufacturing." *Journal of Industrial Technology*, 45(3), 210-225.

- 2 Thompson, R. (2023). "Digital Transformation of Production Processes." *International Journal of Production Research*, 61(4), 1287-1301.
- 3 Zhang, Y., & Liu, Q. (2022). "Machine Learning Applications in Instrumentation Manufacturing." *Journal of Intelligent Manufacturing*, 33(6), 1625-1642.
- 4 Li, X., & Park, J. (2022). "Text Analysis Methods for Manufacturing Documentation." *Journal of Computing and Information Science in Engineering*, 22(3), 031007.
- 5 Lee, H., & Collins, T. (2023). "Semantic Understanding of Technical Documentation." *Applied Artificial Intelligence*, 37(1), 2150007.
- 6 Zhao, Q., & Morgan, A. (2022). "Integrated Text Processing Systems for Production Optimization." *Journal of Manufacturing Technology Management*, 33(9), 1548-1565.
- 7 Wang, B., & Gonzalez, C. (2023). "Large Language Models in Worker Report Analysis." *Artificial Intelligence in Engineering*, 16, 100133.
8. Vladyslav, Y., & Bronnikov, A. (2020, October). ANALYSIS OF THE CMMI MODEL APPLICATION FOR SOLVING THE TASKS OF CPPS CONTROL PROCESSES AUTOMATION DEVELOPMENT. In *The 4 th International scientific and practical conference "Actual trends of modern scientific research"*(October 11-13, 2020) MDPC Publishing, Munich, Germany. 2020. 386 p. (p. 128).
9. Yevsieiev, V. V., & Bronnikov, A. I. (2020). Development of databases interconnection "essences" information model for cyber-physical production systems additive cyber design creation automation. *Збірник Наукових Праць НУК*, №3. С.56-62. DOI [https://doi.org/10.15589/znp2020.3\(481\).7](https://doi.org/10.15589/znp2020.3(481).7)
10. Yevsieiev V., Bronnikov A. Information systems development methodologies application analysis for cyber-physical production systems development. III International scientific-practical conference "Theory, science and practice" (Japan, Tokyo, 5–8 October 2020). P. 398–401. DOI: 10.46299/ISG.2020.II.III
11. Yevsieiev V., Bronnikov A. Analysis of the cyber-physical production systems implementation impact to achieve the goals of lean production. The IIth International scientific and practical conference «Development of scientific and practical approaches in the era of globalization» (USA, Boston, 28–30 September. 2020). P.221–226. DOI:10.46299/ISG.2020.II.II.
12. Attar, H., & et al.. (2022). Zoomorphic Mobile Robot Development for Vertical Movement Based on the Geometrical Family Caterpillar. *Computational Intelligence and Neuroscience*, 2022, Article ID 3046116, <https://doi.org/10.1155/2022/3046116>.
13. Abu-Jassar AT, Attar H, Amer A, et al. Remote Monitoring System of Patient Status in Social IoT Environments Using Amazon Web Services (AWS) Technologies and Smart Health Care. *International Journal of Crowd Science*, 2024, <https://doi.org/10.26599/IJCS.2023.9100019>
14. Abu-Jassar AT, Attar H, Amer A, et al. Development and Investigation of Vision System for a Small-Sized Mobile Humanoid Robot in a Smart Environment. *International Journal of Crowd Science*, 2024, <https://doi.org/10.26599/IJCS.2023.9100018>
15. Chala, O., Yevsieiev, V., Maksymova, S., & Abu-Jassar, A. (2025). MATHEMATICAL MODEL BASED ON MULTI-AGENT REINFORCEMENT LEARNING (MARL) AND PARTIALLY OBSERVABLE MARKOV DECISION PROCESS (POMDP) FOR MODELING CARGO MOVEMENT FOR A MOBILE ROBOTS GROUP. *Multidisciplinary Journal of Science and Technology*, 5(4), 480-489.

Науковий керівник: *Бронніков Артем Ігорович, доц., к.т.н., доцент кафедри КІТАР Харківського національного університету радіоелектроніки.*