

УДК 004.8

Залуцька О.О., Скрипльонок С.Д.

Хмельницький національний університет

ПРОЄКТУВАННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ СЕМАНТИЧНОЇ ОБРОБКИ ТЕКСТІВ ДЛЯ ВИЗНАЧЕННЯ ТОНАЛЬНОСТІ

Подано проєкт інформаційної системи семантичної обробки україномовних коротких повідомлень для визначення тональності, зорієнтований на керованість, відтворюваність і прозорість. Архітектуру побудовано як багатошаровий застосунок з чітким розмежуванням відповідальностей: прийом і буферування даних, підготовка та нормалізація тексту, семантична аналітика з формуванням ознак, агрегування індикаторів тональності, збереження результатів і артефактів аудиту. Логічна модель даних включає сутності документів, токенів, словосполучень, частотних і вагових профілів, лексиконних значень, проміжних агрегатів і фінальних вердиктів, пов'язані з версіями конфігурацій і довідників. Інтерфейси розділено на інтеграційні та експертні з фіксованими контрактами і підтримкою міграцій. Передбачено телеметрію, політики доступу за ролями, незмінювані журнали та механізми відновлення, а також стратегії кешування і матеріалізації для забезпечення продуктивності та масштабованості. Запропонована специфікація створює інженерне підґрунтя для промислового впровадження системи з повним аудитним слідом та керованим життєвим циклом конфігурацій.

The paper presents the design of an information system for semantic processing of short Ukrainian texts aimed at sentiment determination with emphasis on controllability, reproducibility and transparency. The architecture follows a layered application with strict separation of concerns, covering data intake and buffering, text preparation and normalization, semantic analytics for feature formation, aggregation of sentiment indicators, and persistent storage of results and audit artifacts. The logical data model defines entities for documents, tokens, phrases, frequency and weight profiles, lexicon values, intermediate aggregates and final verdicts, all tied to versioned configurations and reference resources. System interfaces are split into integration and expert endpoints with stable contracts and migration support. The design embeds telemetry, role based access control, immutable logs and recovery mechanisms, together with caching and materialization strategies to ensure performance and scalability. The specification provides an engineering foundation for production deployment with a complete audit trail and a governed configuration lifecycle.

Сучасні методи обробки природної мови (NLP) є ключовою технологічною основою для побудови інформаційних систем семантичного аналізу, що здатні автоматично інтерпретувати емоційно-оцінні компоненти тексту [1, 2]. За останнє десятиріччя NLP перейшло від використання традиційних статистичних підходів до глибоких нейронних моделей [3], які оперують контекстними представленнями

мовних одиниць та забезпечують суттєве підвищення точності аналізу тональності, семантичного структурування та класифікації текстових даних [4].

У контексті проєктування інформаційної системи для визначення тональності особливу роль відіграють моделі трансформерної архітектури (BERT, RoBERTa, XLM-R) [5, 6], що дозволяють відтворювати глибокі семантичні залежності та враховувати контекст на рівні всього документа [7]. Застосування таких моделей забезпечує здатність до коректного трактування полісемії, довгострокових залежностей [8], змішаних мовних реєстрів та паралінгвістичних маркерів емоційності [9], що є критично важливим для задач точного визначення тональності природної мови [10, 11].

Окремої уваги потребує специфіка україномовних текстів, які характеризуються високою морфологічною варіативністю, наявністю просторіччя, діалектних елементів та активним розвитком неологізмів у цифрових комунікаціях [12]. Ці фактори створюють додаткові виклики при розробленні систем автоматичної тональної класифікації [13], однак одночасно стимулюють розвиток україномовних лінгвістичних ресурсів та спеціалізованих моделей [14, 15]. Поява українських BERT-подібних моделей і зростання обсягів розмічених корпусів суттєво розширюють можливості побудови високоточних інформаційних систем семантичної обробки текстів [16, 17].

Актуальність розробки інформаційної системи семантичної обробки текстів визначається необхідністю стабільного і відтворюваного аналізу тональності у реалістичних потоках даних [18]. Виробничі сценарії потребують не лише точності на контрольних вибірках, а передусім керованості параметрами, прозорості прийняття рішень і можливості аудиту. Проєктування такої системи передбачає цілісну узгодженість між архітектурою застосунку, схемою зберігання проміжних артефактів, семантичними модулями та методикою тестування. Центральним принципом стає документування кожного кроку обробки у вигляді відтворюваних артефактів, завдяки чому будь-який результат може бути пояснений і відтворений за ідентичних умов.

Метою проєктування є створення керованої, відтворюваної та прозорої системи, здатної приймати короткі україномовні повідомлення, стабілізувати їх уніфікованим поданням, формувати семантичні ознаки та обчислювати тональність із фіксацією усіх проміжних станів. Вимоги формуються як інженерні обмеження до архітектури, моделі даних, інтерфейсів і експлуатаційних властивостей. Ключовими принципами є розмежування відповідальностей, версіонування конфігурацій і словників, матеріалізація артефактів для аудиту, відсутність прихованих побічних ефектів і можливість детермінованого повторного прогону за часовою міткою.

Архітектурну основу становить багатoshаровий застосунок із чіткою ізоляцією контурів. На зовнішньому периметрі розміщується підсистема прийому, що працює з потоками коротких текстів і забезпечує нормалізоване, безпечне буферування. Далі діє шар підготовки, задачі якого включають уніфікацію графематики, стабілізацію токенизації та побудову службових метаданих для подальшого трасування. Шар семантичної аналітики відповідає за формування ознак та обчислення індикаторів тональності за узгодженими контрактами. Шар прийняття рішень агрегує сигнали в остаточне подання тональності та передає результат у підсистему збереження. Паралельно працює контур спостережності, що накопичує журнали параметрів, артефакти проміжних обчислень і довідники з їхніми версіями. Така композиція забезпечує локалізацію змін і дозволяє модернізувати окремі модулі без каскадного впливу на весь конвеєр.

Модель даних проектується як логічно зв'язана схема, орієнтована на прозорість і повторюваність. Передбачено сутності для документів, токенів, словосполучень, профілів частот, вагових оцінок ознак, лексиконних значень, проміжних агрегатів і фінальних вердиктів. Кожний запис пов'язується з конфігураційною сесією, що має часову мітку, і з набором версій довідників. Залежності організовано таким чином, щоб важкі у перерахунку об'єкти могли бути матеріалізовані та повторно використані. Це зменшує латентність і дає можливість виконувати локальні перебудови, зберігаючи стабільність загальної поведінки. Схема підтримує повний шлях пояснюваності, оскільки у ній фіксуються як вихідні повідомлення, так і їхні нормалізовані представлення, ключові одиниці, значення ваг, параметри порогів і остаточні рішення. Схема таблиць бази даних наведена на рисунку 1.



Рисунок 1 – Схема бази даних

Схема бази даних спроектована як нормалізована предметна модель для зберігання первинних текстових даних, морфологічних одиниць, їхніх поєднань і позицій у повідомленнях, а також проміжних і підсумкових артефактів семантичного аналізу. Логічний центр становить таблиця Texts, що акумулює короткі цифрові повідомлення у вигляді ідентифікованих записів із назвою, повним текстом, часовою міткою додання та агрегованими характеристиками довжини за словами і символами. Ця сутність є джерелом зв'язків для всіх подальших артефактів аналізу, забезпечуючи відтворюваність на рівні конкретного документа.

Лексична інформація представлена таблицею Words, де кожне слово зберігається як незалежна семантична одиниця з текстовою формою та довжиною. Посилання на Infinitives та FormTypes дозволяють відокремити лемму й тип словоформи від поверхневої форми, що усуває надлишковість і фіксує морфологічну норму для подальших зіставлень. Модель підтримує сценарії, коли кілька різних словоформ відносяться до одного інфінітива, а також розрізняє граматичні варіанти через окремий довідник типів форм.

Фразеологічний рівень подано сутністю WordCombs, яка описує словосполучення як окремі інформаційні носії з власним текстовим представленням, довжиною та лічильником зустрічей. Зв'язок між словосполученням і словами, що його утворюють, реалізовано через проміжну таблицю WordsInCombinations: тут для кожної комбінації фіксуються складові слова і порядок їх розташування, що дає можливість відтворити структурну форму та відокремити її від конкретних появ у документах. Така декомпозиція підтримує канонічне збереження шаблону словосполучення та унеможливорює дублювання даних.

Позиційні відомості зберігаються двома взаємодоповнювальними сутностями. Таблиця WordApperance (у подальшому доцільно уніфікувати написання як WordAppearance) відображає появи слів у межах конкретного повідомлення, фіксуючи посилання на слово та текст разом із позиціями за словами і символами. Аналогічно, WordCombApperance (WordCombAppearance) відстежує входження словосполучень у текстах із прив'язкою до абсолютних позицій. Розмежування цих двох рівнів дозволяє, з одного боку, реконструювати точну послідовність лексем у повідомленні, а з іншого — аналізувати сталі фрагменти як окремі ознаки, не втрачаючи їх прив'язки до первинних даних.

Шар семантичних обчислень реалізовано через спеціалізовані сутності для ключових одиниць. У KeyWordsInMessages зберігаються зв'язки між текстом і словом, що набуло статусу ключового, разом із чисельними показниками, отриманими на етапі векторизації. Передбачено поля для значень TF-IDF та порогових критеріїв відсікання за різними методами й різними шкалами позиціонування. Схема симетрично підтримує параметри для символорієнтованого

та словоорієнтованого підрахунку, що дає змогу відтворювати умови експерименту й проводити регресійні порівняння. Для багатослівних ознак аналогічні дані акумулює KeyWordCombinationsInTexts, де ключовою одиницею виступає словосполучення, пов'язане з відповідним текстом і доповнене набором порогових атрибутів для застосованих методів відбору. Вказані сутності виконують роль «аудитних таблиць» векторизації: вони не лише тримають підсумкові значення, а й відображають конфігураційний контекст ухвалення рішень.

Реляційна структура дотримується принципів третьої нормальної форми: текст, слово, словосполучення та їхні появи розведені у відокремлені сутності; довідникова інформація про лема та типи форм винесена за межі операційних таблиць; числові метрики зберігаються поруч із ключами, що визначають їхню область чинності. Первинні ключі забезпечують унікальність записів, тоді як зовнішні ключі фіксують цілісність посилань між рівнями лінгвістичної абстракції й конкретними документами. Для забезпечення продуктивності індекси доцільно створити на полях зв'язків FK_text, FK_word, FK_wordComb, а також на текстових полях wordContent та wCombContent з урахуванням налаштувань сортування та мовної локалі. Для позиційних таблиць доречно встановити унікальні обмеження на композиції (FK_text, wordPosition) і (FK_text, sympolPosition), що унеможливило дублікати та гарантує однозначність координат у межах повідомлення.

Семантика часової прив'язки реалізується через поле дати й часу у Texts, яке визначає базову мітку для відтворення прогонів. Усі похідні артефакти можуть бути додатково з'єднані з конфігураційними версіями й словниковими релізами через відповідні довідники, що забезпечує повний аудит параметрів обчислень. Узгодженість імен і типів даних підтримує детермінованість обмінних контрактів між прикладними модулями, а розділення на словникові, позиційні та метрик-орієнтовані сутності знижує зв'язаність і полегшує еволюцію схеми без порушення цілісності історичних даних.

Інтерфейси системи поділяються на інтеграційні та експертні. Інтеграційні інтерфейси приймають повідомлення з зовнішніх сервісів у стандартизованому форматі та повертають структуровану відповідь із тональністю і базовими метаданими. Експертні інтерфейси забезпечують доступ до трас рішень, включаючи перелік ключових одиниць з їхніми вагами, застосовані конфігурації та пов'язані журнали. Контракти між шарами визначаються як стабільні схеми обміну, що унеможливають неявні зміни структури даних. Внутрішні API фіксують версії, підтримують міграції та валідацію вхідних параметрів, щоб запобігати розходженням між модулями під час оновлень.

Нижче наведено спроектовану діаграму класів для системи визначення тональності (рисунок 2).

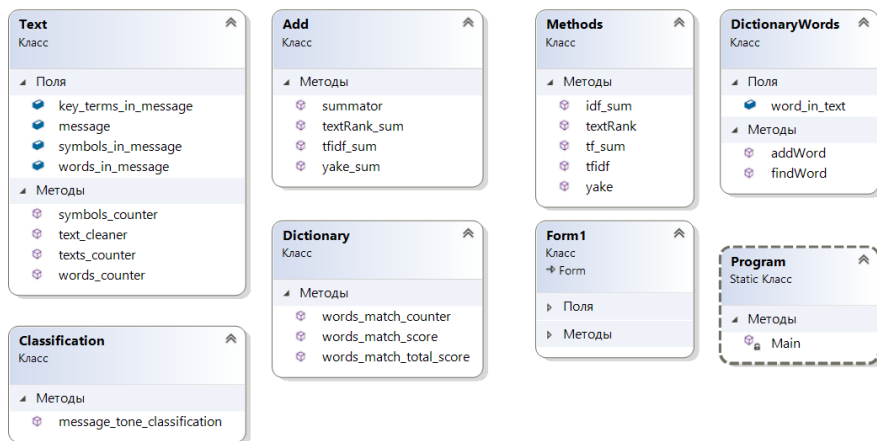


Рисунок 2 – Діаграма класів

Робота ініціюється з точки входу класу Program, який відповідає за коректне завантаження конфігурацій, підготовку середовища виконання та запуск основної форми користувацького інтерфейсу. Графічний інтерфейс реалізовано у вигляді форми Form1, що виконує роль центрального вузла взаємодії з користувачем. Форму організовано навколо тематичних вкладок, які охоплюють повний цикл опрацювання коротких україномовних повідомлень: від первинного завантаження і стабілізації тексту до семантичної інтерпретації та узагальнення результатів. Така композиція забезпечує логічну послідовність дій, мінімізує контекстні перемикання та робить процедури обробки відтворюваними.

Вкладка, присвячена роботі з базою повідомлень, уможливорює імпорт, попередній перегляд і добір матеріалу для аналізу з фіксацією метаданих. Тут відбувається підготовка корпусу й перевірка технічної придатності записів до подальшого опрацювання. Вкладка появ семантичних одиниць формує для вибраних текстів їхнє стабілізоване подання, фіксує токенізацію, частотні профілі та вектори ключових одиниць і надає користувачеві інспектовані проміжні результати. На вкладці аналізу повідомлень здійснюється зіставлення із словниковими ресурсами, агрегування сигналів і формування підсумкової класифікації з одночасним збереженням трас пояснюваності. Кожен елемент інтерфейсу пов'язаний із конкретним шаром логіки, що гарантує однозначну відповідність між дією користувача і викликаними сервісами.

Функції первинної обробки тексту інкапсульовано в окремому класі Text, який утворює стабільне підґрунтя для подальшої семантичної обробки. Методи класу виконують нормалізацію шляхом усунення цифрових і службових символів,

уніфікації реєстру та очищення невидимих артефактів кодування, що могли б викривляти частотний профіль. Окремі процедури відповідають за підрахунок кількості доступних повідомлень, слів і символів, завдяки чому формується базова статистика для контролю якості та налаштування порогів допуску. Сукупність цих операцій забезпечує інваріантність подальшої токенізації та робить результати наступних шарів порівнянними між різними джерелами даних.

Видобування інформативних ознак реалізовано в класі `Methods`, де зосереджено алгоритмічні засоби для побудови векторних представлень. Локальна помітність терміна визначається через частотні характеристики, тоді як глобальна розрізнявальна здатність регулюється інверсним зважуванням, що знижує вплив повсюдних лексем. Паралельно застосовуються безвчительні процедури, орієнтовані на внутрішньодокументні сигнали, зокрема механізми, які оцінюють позиційні ефекти, контекстну різноманітність та графову зв'язність кандидатів. Після узгодження шкал отримується композитний вектор ключових одиниць, придатний як для лексиконного зіставлення, так і для подачі на класифікаційні модулі. Такий дизайн забезпечує баланс між корпусною дисципліною і чутливістю до риторичної структури коротких повідомлень.

Робота зі словниками інтерпретації тональності зосереджена у класі `Dictionary`, який відповідає за операції зіставлення та агрегування оцінок. Спочатку визначаються перетини між відібраними ключовими одиницями і записами словника з урахуванням нормалізованих форм та можливих морфологічних варіантів. Далі кожній знайденій відповідності надається вагове значення згідно з валентною шкалою, після чого підрахунок інтегрується у зважений підсумок для всього повідомлення. Клас зберігає проміжні внески на рівні окремих термінів, що формує прозорі підстави для подальшої інтерпретації та спрощує аналіз джерел впливу на підсумковий знак.

Фінальне узгодження ознак у тональний вердикт виконує клас `Classification`, який здійснює перетворення підсумкової оцінки на категоріальне рішення та супровідні ймовірнісні показники. Перетворення враховує нейтральну зону з фіксованими порогамі, що запобігає поспішним рішенням у прикордонних випадках і підвищує стабільність поведінки на неоднорідних потоках. Результат подається у вигляді структурованого висновку, який містить категорію тональності, інтенсивність сигналу, посилання на використані версії словників і параметрів та скорочену трасу пояснюваності з переліком провідних семантичних одиниць.

Міжкласові взаємодії опрацьовано через чіткі контрактні інтерфейси. `Form1` ініціює операції класу `Text` для підготовки даних, передає стабілізовані тексти у `Methods` для формування ознак, скеровує отримані профілі до `Dictionary` для зіставлення зі словниками та, нарешті, викликає `Classification` для одержання підсумкового рішення. Кожний крок супроводжується фіксацією параметрів

виконання і часових міток, що уможливило аудит і відтворюваність. Така послідовність забезпечує низьку зв'язаність між підсистемами, високу когерентність усередині класів і локалізованість змін під час модернізації окремих алгоритмічних компонентів.

Експлуатаційні вимоги враховують продуктивність, масштабованість і стійкість. Для зменшення затримок застосовується кешування результатів підготовки і семантичних обчислень, а також матеріалізація часто використовуваних представлень. Обробка організується пакетами з лінійною залежністю часу від обсягу вхідних даних. Стратегія масштабування передбачає горизонтальне рознесення підсистем за функціональними ролями, що дозволяє окремо збільшувати ресурси підготовки, семантики або прийняття рішень залежно від профілю навантаження. Усі компоненти проєктуються з урахуванням відмовостійкості, передбачено механізми повторних спроб, ідемпотентність обробки та консистентні транзакції для критичних операцій запису.

Безпека і відповідність етичним вимогам закладаються на рівні проєктних рішень. Система працює з анонімізованими або агрегованими представленнями, мінімізує обсяг персональних даних і розділяє доступ до технічних і змістових артефактів. Політика контролю доступу визначається ролями та підтримує принцип мінімальних привілеїв. Усі експортовані результати супроводжуються метаданими про версії конфігурацій і часові мітки, що забезпечує можливість зовнішнього аудиту. Журнали доступу та зміни параметрів зберігаються у незмінюваному вигляді, що гарантує довіру до відтворюваності.

Управління конфігураціями та довідниками є окремим аспектом проєктування. Кожна зміна порогів, словників, правил нормалізації або форматів контрактів оформлюється як версійована подія із вказанням причини, відповідального і кола впливу. Підтримується можливість повернення до попереднього стану за однією командою. Для критичних оновлень застосовується подвійний запуск у паралельному контурі, що дає змогу порівнювати виходи різних версій на однакових вхідних потоках до моменту повного переключення.

Спостережність і керуваність забезпечуються єдиним шаром телеметрії. У ньому акумулюються технічні метрики затримок та використання ресурсів, статистика по класах і зведення щодо відсотка повідомлень у нейтральній зоні. Артефакти трас зберігаються з прив'язкою до ключів сесії, що дозволяє реконструювати повний шлях рішення для будь-якого повідомлення. Передбачено механізм сигналів про деградацію, який ініціює захисні сценарії, наприклад автоматичне розширення нейтральної зони або переключення на попередню стабільну конфігурацію до завершення аналізу причин.

Розгортання системи планується у конфігурації з поділом середовищ. Окреме середовище використовується для розробки та інтеграційного узгодження

контрактів, окреме для передпромислової валідації з доступом до повних трас і незмінюваних журналів, а окреме для промислової експлуатації з обмеженим доступом і жорсткими політиками версіювання. Міграції схеми даних проводяться за сценаріями, що не потребують зупинки сервісу, а всі зміни контрактів супроводжуються автоматичною перевіркою сумісності.

Підсумовуючи, проектування інформаційної системи семантичної обробки текстів для визначення тональності зосереджується на архітектурній ясності, відтворюваності і керованості. Узгоджена багатоваріантова структура, прозора модель даних із матеріалізованими артефактами, стабільні інтерфейси, дисципліна версіювання і вбудована спостережність формують інженерне підґрунтя для подальшого масштабованого впровадження без компромісів щодо прозорості прийняття рішень і можливості аудиту.

Перелік посилань

1. Siino M., Falco M., Croce D., Rosso P. Exploring LLMs Applications in Law: A Literature Review on Current Legal NLP Approaches. IEEE Access. 2025.
2. Akter Semi M. M., Uddin M. B., Sultana S., Tamanna M., Uddin A., Ahmed K. R. AI-Driven Education: Integrating Machine Learning and NLP to Transform Child Learning Systems. International Journal of Advanced Computer Science & Applications. 2025. Vol. 16, No. 6.
3. Yu J. H., Chauhan D. Trends in NLP for Personalized Learning: LDA and Sentiment Analysis Insights. Education and Information Technologies. 2025. Vol. 30, No. 4. Pp. 4307–4348.
4. Poddar S., Koley P., Misra J., Ganguly N., Ghosh S. Towards Sustainable NLP: Insights from Benchmarking Inference Energy in Large Language Models. Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies. April 2025. Pp. 12688–12704.
5. Timoneda J. C., Vera S. V. BERT, RoBERTa, or DeBERTa? Comparing Performance Across Transformers Models in Political Science Text. The Journal of Politics. 2025. Vol. 87, No. 1. Pp. 347–364.
6. Shevchuk P., Molchanova M., Mazurets O. Software for Text Messages Reliability Analysis Based on the Machine Learning Models Ensemble. Proceedings of IV International Scientific and Practical Conference «Innovative research and perspectives of the development of science and technology». January 29-31, 2024. Stockholm, Sweden. 2024. Pp. 347-354.
7. Mazurets O., Sobko O., Vit R., Pasternak V. Practical Approach for Detection by Deep Learning of Target Objects of Subject Area Based on Semantic Connectivity Indicators in Audio Database. Proceedings of XXIV International Scientific and Practical Conference «Modern Scientific Challenges are the Driving Force of the Development of Scientific Research». May 22-24, 2024. Bruges, Belgium. International Scientific Unity. 2024. Pp. 91-96.
8. Молчанова М.О., Мазурець О.В., Собко О.В., Кліменко В.І., Андрощук В.І. Метод неймережевого виявлення кібербулінгу з використанням хмарних сервісів та об'єктно-орієнтованої моделі. Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки. Хмельницький, 2024. №2 (333). С. 200-206.

9. Mazurets O.V., Sobko O.V., Molchanova M.O., Zalutska O.O., Yurchak A.V. Practical Implementation of Neural Network Method for Stress Features Detection by Social Internet Networks Posts. *Global Science: Prospects and Innovations. Proceedings of the II International Scientific and Theoretical Conference «Scientific Review of the Actual Events, Achievements and Problems»*. May 31, 2024. Berlin, Federal Republic of Germany: International Center of Scientific Research. 2024. Pp. 160-167.
10. Sobko O., Mazurets O., Didur V., Chervonchuk I. Recurrent Neural Network Model Architecture for Detecting a Tendency to Atypical Behavior Of Individuals by Text Posts. *Theoretical and Practical Aspects of Modern Research. Proceedings of XXVI International scientific and practical conference*. June 5-7, 2024. International Scientific Unity. Ottawa, Canada. 2024. Pp. 113-117.
11. Mazurets O., Molchanova M., Klimentenko V., Prosvitliuk M Practice Implementation of Neural Network Model BART-Large-CNN for Text Annotation. *Prospects of Scientific Research in the Conditions of the Modern World. Proceedings of XXVII International scientific and practical conference*. June 12-14, 2024. Rotterdam, Netherlands. 2024. Pp. 97-102.
12. Blazhuk V., Mazurets O., Zalutska O. An Approach to Using the mBERT Deep Learning Neural Network Model for Identifying Emotional Components and Communication Intentions. *The Impact of Scientific Research on the Development of the Modern World. Proceedings of the XLIV International scientific and practical conference*. October 23-25, 2024. Dubrovnik, Croatia. 2024. Pp. 79-84.
13. Tymofiiiev I., Mazurets O., Hardysh D., Molchanova M. Neural Network Dual Architecture for Depression Detection Using Cloud Services. *Scientific Research in the Era of Digital Technologies: Challenges and Opportunities. Proceedings of the XLVI International scientific and practical conference*. November 6-8, 2024. Barcelona, Spain. 2024. Pp. 84-88.
14. Yurchenko D., Mazurets O., Didur V., Molchanova M. Approach to Using Cloud Services for Visual Analytics of Neural Network Analysis of Texts Emotional Tonality. *The Future of Scientific Discoveries: New Trends and Technologies. Proceedings of the XLVII International scientific and practical conference*. November 13-15, 2024. Marseille, France. 2024. Pp. 108-113.
15. Mazurets O., Tymofiiiev I., Dydo R. медвApproach for Using Neural Network BERT-GPT2 Dual Transformer Architecture for Detecting Persons Depressive State. *Ricerche scientifiche e metodi della loro realizzazione: esperienza mondiale e realtà domestiche. Raccolta di articoli scientifici con gli atti della VI Conferenza scientifica e pratica internazionale*. 15 novembre, 2024. Bologna, Repubblica Italiana. 2024. Pp. 147-151.
16. Hladun O., Mazurets O., Molchanova M., Sobko O. Real Time Detection the Person Emotion State Using Neural Network. *Scientific Research: Modern Innovations and Future Perspectives. Proceedings of the 2 International scientific and practical conference*. November 25-27, 2024. Montreal, Canada. 2024. Pp. 119-123.
17. Mazurets O., Vit R. Practical Application of Method of Thematic Classification of Text Information Using LDA. *Information Technology and Implementation (Satellite). Proceedings 11th International Conference*. November 21, 2024. Kyiv, Ukraine. 2024. Pp. 151-152.
18. Murava V., Zalutska O., Didur V., Mazurets O. Software architecture of information system for exchanging LLM thematic prompts. *Global Trends in the Development of Information Technology and Science. Proceedings IV International Scientific and Practical Conference*. June 25-27, 2025. Stockholm, Sweden. Pp. 121-127.