

УДК 004.8

Масловська В.В., Залуцька О.О.

Хмельницький національний університет

## ОСОБЛИВОСТІ РОЗРОБКИ ТА ТЕСТУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ ВИЗНАЧЕННЯ ТОНАЛЬНОСТІ В УКРАЇНОМОВНИХ ПОВІДОМЛЕННЯХ

*У роботі висвітлено особливості розробки та тестування інтелектуальної системи визначення тональності в україномовних повідомленнях. Показано вибір архітектури на основі поєднання нормалізації тексту, безвчительного виокремлення ключових одиниць та лексиконної валентності з числовою шкалою. Обґрунтовано використання середовища .NET і MS SQL Server для збереження проміжних артефактів і журналів параметрів, що забезпечує відтворюваність і аудит. Описано протоколи випробувань із модульними перевітками, перехресною валідацією та аналізом хибних рішень, а також регресійні серії для оцінювання чутливості порогів. Зафіксовано стабілізацію метрик за рахунок розширеної нейтральної зони та контрольований вплив заперечення і інтенсифікаторів.*

*The study presents the design and testing specifics of an intelligent system for sentiment detection in Ukrainian short messages. The architecture integrates text normalization, unsupervised key unit extraction, and a lexicon based valence model with a numeric scale. The system is implemented in .NET with MS SQL Server to persist intermediate artifacts and parameter logs, which supports reproducibility and audit. The testing protocol covers module level checks, cross validation, analysis of erroneous outcomes, and regression series to assess threshold sensitivity. The results indicate steadier metrics due to an expanded neutrality zone and a controlled treatment of negation and intensifiers.*

Останнім часом автоматизоване визначення тональності україномовних коротких повідомлень виходить за межі лабораторних прототипів і переходить до виробничих сценаріїв, де на перший план виходять прозорість прийняття рішень, відтворюваність обчислень і керуваність параметрами [1]. Короткий формат повідомлень, висока стилістична мінливість, перемишування розмовних конструкцій із книжними, домішки емодзі, навмисні орфографічні варіації та фрагментарний синтаксис ускладнюють застосування моделей, побудованих на глибоких контекстних репрезентаціях без належної доменної адаптації [2, 3]. За таких умов практично доцільним є підхід, що тримає баланс між інженерною дисципліною та інтерпретованістю [4]: попередній контроль якості корпусу, безвчительне виділення ключових одиниць і лексиконна валентність із числовою шкалою, яка дозволяє пояснити внесок кожного слова у підсумкову полярність [5].

Методи обробки природної мови є ключовими технологіями сучасних інтелектуальних систем, що працюють із текстовими даними [6, 7]. Завдяки переходу від традиційних статистичних методів [8] до нейромережевих моделей [9, 10], зокрема трансформерних архітектур [11], NLP забезпечує високий рівень точності в задачах класифікації [12], тонального аналізу [13] та семантичної інтерпретації текстів [14]. Це робить такі методи особливо цінними для розроблення систем автоматичного визначення емоційної забарвленості повідомлень, що є критично важливою складовою сучасних інформаційних сервісів, чат-ботів, систем модерації контенту, моніторингу соцмереж та інструментів аналітики громадської думки [15, 16, 17].

У контексті україномовних текстів застосування NLP стикається з низкою специфічних викликів, пов'язаних із недостатнім обсягом якісних розмічених датасетів, варіативністю української морфології та обмеженою кількістю спеціалізованих мовних моделей [18]. Попри це, останні роки характеризуються зростанням числа україномовних корпусів, появою BERT-подібних моделей української мови та активним розвитком інструментів лематизації, токенизації та частини мовного аналізу. Це створює сприятливі умови для побудови високоточних систем визначення тональності [19].

Створення системи розпочиналося з формулювання вимог до вхідних даних. Було прийнято, що об'єктом обробки є короткі україномовні повідомлення, нерідко з мінімальним контекстом і локальними тригерами, де емоційний знак часто задається одним-двома словами або стійким словосполученням. Нормалізація була спроектована як обов'язковий початковий крок, але з чіткою межею її повноважень: уніфікація лапок і тире, стабілізація апострофів, усунення невидимих службових символів, упорядкування пробілів і перетворення емодзі у нейтральні маркери, сумісні з подальшою токенизацією, без спроби повного лематизування, яке могло б привнести невинуватому неоднозначність. Такий дизайн нормалізації не намагається «зробити все одразу», але забезпечує стабільні умови для частотного аналізу й побудови ключових одиниць. Схему підсистем наведено на рисунку 1.

Наступним структуро-утворювальним етапом був контроль якості корпусу до початку тонального аналізу. Показник «нудоти» трактувався як індикатор неприродної повторюваності, що в коротких текстах здатна штучно зрушувати полярність до крайніх значень. Паралельно перевірявся профіль розподілу частот: очікувана квазілінійність у логарифмічних координатах вважалася сигналом природності, тоді як виразні «полиці» або розриви свідчили про спотворення даних. На підставі цих перевірок приймалися процедурні рішення про допуск або очищення текстів. Принциповим було документувати кожний поріг і кожну причину відсікання у спеціальних таблицях контролю якості, щоби будь-який прогін згодом можна було відтворити з ідентичними вхідними умовами.

Після допуску корпусу активувався беззвчительний модуль виділення ключових одиниць. Його роль полягала у локалізації смислових центрів короткого повідомлення без звернення до тональних індикаторів. Поєднувалися дві

взаємодоповнювальні процедури: TF-IDF як міра корпусної специфічності слів і стійких словосполучень та TextRank як графова рангово-переходова модель, що чутлива до локальних зв'язків у вікні співзвучіччя. Обидва сигнали нормувалися в єдине числове поле, після чого виконувалося злиття з пріоритетом стабільності на малих вибірках. У надкоротких повідомленнях більшу довіру надавали ранговому сигналу, чутливому до структурної ролі лексем у висловлюванні, тоді як зі зростанням корпусу ваги TF-IDF і TextRank вирівнювалися. Важливо, що цей модуль не виконував класифікацію сам по собі; його завданням було сформулювати вагову опору для наступного контуру, який приймає рішення про полярність.

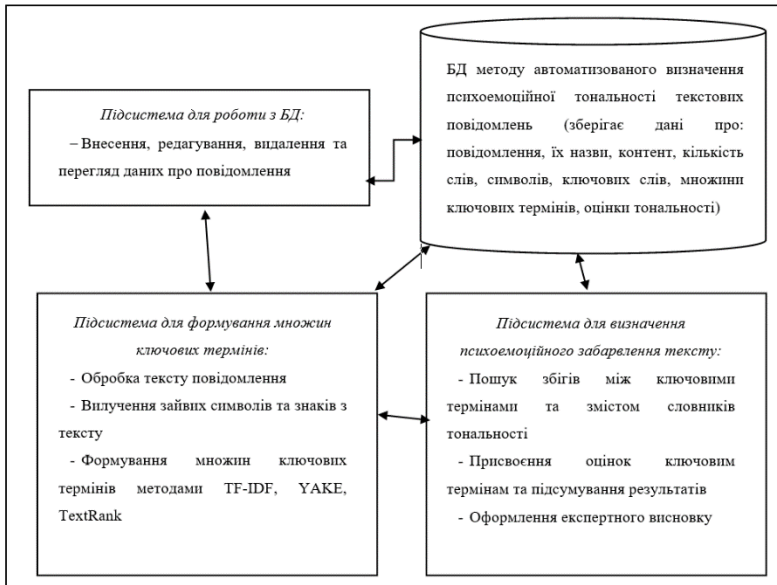


Рисунок 1 – Схема підсистем інтелектуальної системи

Ядро системи становить лексиконна валентність, де кожній лексемі надано числову оцінку у шкалі від одиниці до дев'яти. Числова природа лексикону дає можливість будувати пояснювані рішення, у яких видно, як саме окремі лексеми впливають на сумарну полярність. Агрегування реалізовувалося як зважене середнє, де ваги походили з беззвичайних моделей ключових одиниць, а власне числові значення — із лексикону. Окрему увагу приділено локальній логіці заперечення: у вузькому вікні навколо дієслівної групи або прикметникового предикатива відбувалося інвертування або зменшення ефекту вихідної валентності, що дозволяло уникати помилкових позитивів, спричинених механічною сумою позитивних лексем. Інтенсиватори та димінутиви трактувалися як модифікатори відхилення від нейтральної зони з верхньою межею впливу, щоб не допустити

домінування стилістичного маркера над змістовою частиною висловлювання. Кінцеве рішення приймалося за трипороговою схемою з фіксованою «мертвою зоною» навколо нуля, призначеною для відсікань випадкових рішень у прикордонних ситуаціях. Параметри порогів і версії лексикону фіксувалися в конфігураційних таблицях бази даних разом із часовою міткою, що створювало прозору історію змін і спрощувало аудит.

Архітектурно система реалізована у середовищі .NET із використанням MS SQL Server як центральної СКБД. Розширована побудова включала сервіс прийому й нормалізації, аналітичний шар беззвичайних обчислень і лексиконного агрегування, підсистему прийняття рішень та шар збереження проміжних артефактів. Структура бази даних містила таблиці документів, токенів, словосполучень, частотних профілів, ваг TF-IDF і TextRank, лексиконних значень, проміжних агрегатів і фінальних вердиктів, а також журнали параметрів і версій. Такий устрій уможлилював не тільки перевіряти будь-який крок постфактум, а й вибірково переобчислювати окремі шари без повного перезапуску конвеєра, що важливо для масштабування і для серійних експериментів у процесі тестування.

Методика тестування була побудована на двох рівнях. Модульний рівень перевіряв коректність і стабільність кожного етапу окремо. Для нормалізації оцінювалася інваріантність токенизації до графічних варіацій і невидимих символів; вимірювалося, наскільки різні джерела з різною пунктуаційною дисципліною дають однаковий результат. Для скринінгу якості оцінювалася чутливість показника «нудоти» до довжини і жанру, перевірялася стійкість цпфівського профілю в тематично рознесених підвибірках. Беззвичайний модуль ключових одиниць тестувався на узгодженість із людською інтуїцією: для вибраних підмножин експерти попередньо позначали смислові ядра, після чого порівнювали ранжування системи та ручні референси. В лексиконному контурі проводилися контрольні серії на шаблонах заперечення та інтенсиваторах, щоб переконатися у відсутності систематичних перекосів до крайнього позитиву чи негативу в умовах короткого контексту.

Інтегральний рівень тестування оцінював поведінку системи як цілого. Проводилася перехресна перевірка з поділом на незалежні підвибірки; для кожної з трьох категорій: позитивної, нейтральної та негативної – обчислювалися точність, повнота і  $F_1$ -міра, а також макро- і зважені агрегати, чутливі до дисбалансу класів, який у коротких повідомленнях виникає майже завжди. Кожна серія супроводжувалася якісним розбором помилок. У хибно позитивних рішеннях домінували іронічні патерни, інверсії сенсу за рахунок контексту поза повідомленням і поєднання позитивних емоцій з негативним вербальним змістом; у хибно негативних зауважували оцінки з м'якими інтенсифікаторами і складними заперечними ланцюжками, що виходять за межі локального вікна. Каталог помилок зберігався як наростаюча база прикладів із фіксацією тексту, ключових одиниць, вагового внеску кожної лексеми та фінального рішення. На підставі цього каталогу виконувалися точкові корекції: звужувалася або розширювалася «мертва зона»,

уточнювалися пороги для переходу між категоріями, переглядалися правила локального врахування заперечення.

Практично уся інтелектуальна система була покрита тест-кейсами. Один з яких наведено в таблиці 1.

Таблиця 1 – Тест-кейс 1

|  |                     |   |
|--|---------------------|---|
| <b>Тест-кейс ID:</b> АК0004  | <b>Пріоритет:</b> 1 | <b>Створено:</b> 12.05.2025,<br>Залуцька О.   |
| <b>Назва:</b> Перевірка коректного визначення тональності<br><b>Вхідні дані:</b> Текстове повідомлення .txt  |                     |   |
| <b>Кроки</b>   |                     | <b>Очікуваний результат</b>   |
| <p><i>Передумова:</i> на вкладці «База повідомлень» користувач має обрати рядок текстового повідомлення, подальший аналіз якого відбуватиметься. Повідомлення містить явно негативний характер, що необхідно довести.</p> <ol style="list-style-type: none"> <li>Запустити додаток та обрати вкладку «База повідомлень»</li> <li>Обрати рядок текстового повідомлення з негативним емоційним забарвленням.</li> <li>Натиснути кнопку «Аналіз»</li> <li>Порівняти фактичний результат з очікуваним</li> </ol> |                     | <p>Система має довести, що повідомлення містить негативне забарвлення із вірогідністю, більшою за 80%.</p> <p>За результатами аналізу повідомлення із вірогідністю 97.4% містить негативне забарвлення.</p> |
| <b>Результат виконання тест-кейсу:</b> пройдено успішно  |                     |   |

Приклад виконання тесту, наведеному в таблиці 1 для визначення психоемоційного забарвлення наведено на рисунку 2.

Під час серійних експериментів особливий акцент робився на регресійних перевірках, де змінювався лише один клас параметрів. Зокрема, варіювалася ширина вікна для заперечення, після чого оцінювалися метрики і якісні помилки. Було показано, що надто широке вікно провокує помилкові інверсії у випадках, де заперечення належить до іншого синтаксичного вузла, тоді як надто вузьке позбавляє систему здатності вчасно інвертувати полярність у розтягнутих конструкціях. Ця емпірична крива відгуку слугувала аргументом для консервативних налаштувань і була зафіксована в журнальних таблицях параметрів як еталон для наступних змін. Приклад успішного виконання юніт-тестів наведено на рисунку 3.

У питаннях продуктивності було важливо відокремити важкі та легкі шари обчислень. Матеріалізовані представлення для частотних запитів і кеші для співзвуччій дозволили зменшити латентність без модифікації логіки прийняття рішень. Збереження проміжних артефактів збільшує обсяг сховища, але дає змогу вибірково переоцінювати тільки ті компоненти, яких торкнулися зміни параметрів. Перевизначення порогів не потребує повторного підрахунку TF-IDF і TextRank; оновлення словникових норм торкається лише підмножини документів, у яких зустрічаються змінні лексеми. За такої організації даних час обробки масштабується лінійно з розміром партії, а прогнозованість затримок дозволяє планувати ресурси без завищених запасів.

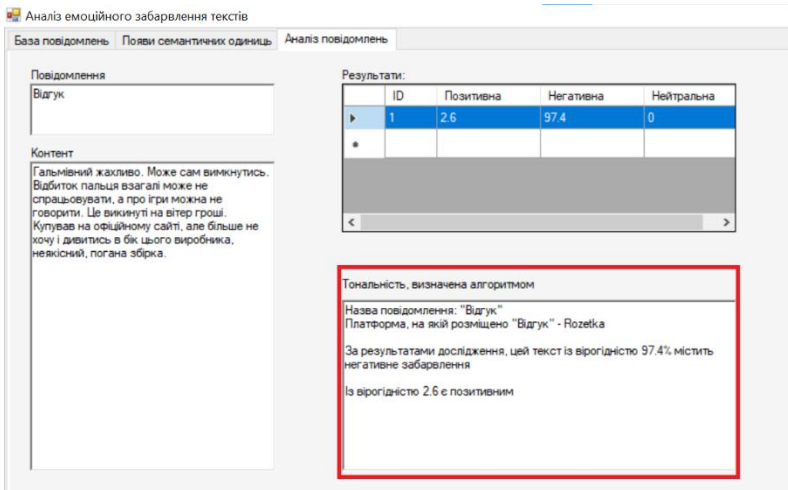


Рисунок 2 – Приклад виконання модульного тесту

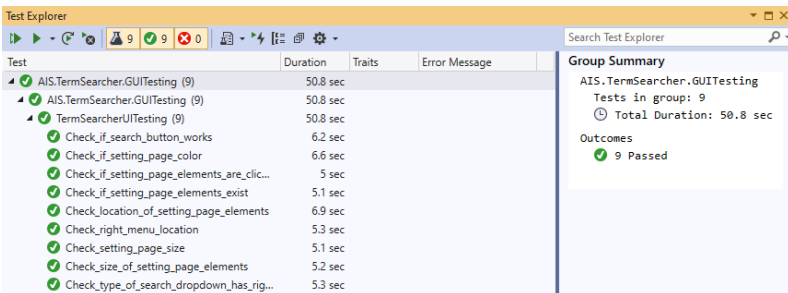


Рисунок 3 – Виконання юніт-тестів

Окремою темою стала інтерпретація нейтральної категорії. В умовах коротких повідомлень нейтральність не є тривіальною серединою між позитивом і негативом. Вона означає недостатність доказів на користь певного полюсу за наявних вагових сигналів. У зв'язку з цим нейтральна зона була розширена і підкріплена регламентом, що забороняє приймати рішучі вердикти на підставі одиничних емоційних маркерів без підтримки ключових словосполучень. Така політика зменшує кількість випадкових коливань і покращує зважені метрики на різномірних потоках.

У межах практичних сценаріїв особливої уваги вимагали повідомлення з високою насиченістю емоції. Щоби уникнути ситуацій, коли піктографічні маркери повністю визначають результат, було запроваджено вагову скидку для емоції та обмеження максимальної кількості ключових одиниць, що бере участь в агрегуванні у надкоротких повідомленнях. Це рішення продемонструвало помітне зниження частки хибно позитивних вердиктів у саркастичних конструкціях із позитивними піктограмами.

Впровадження системи в середовищах із різними тематичними доменами не вимагало змін у конвеєрі. Адаптація зводилася до перегляду порогів допуску корпусу, часткового оновлення словникових норм для специфічних лексем та регресійних перевірок на «заморожених» наборах складних випадків. Версіонування параметрів і словників, пов'язане з часовими мітками, дозволяло повернутися до попереднього стану в разі погіршення метрик. Зауважимо, що запроваджена культура тестування передбачала підтримку стабільного пулу контрольних прикладів з іронією, полісемією і складним запереченням, через які максимально швидко виявлялися деградації після найменших змін у правилах або порогах.

Обмеження підходу не приховувалися й служили орієнтирами для коректної інтерпретації результатів. Лексиконна природа моделі означає залежність від покриття словника та інерційність щодо нових мовних мод і сленгових утворень. Перевірка ціпфівського профілю на дуже малих підвибірках має скринінговий характер і може давати хибні сигнали на жанрових матеріалах із навмисною ритмікою. Іронія і сарказм лишаються найскладнішими випадками, у яких відсутність розгорнутого контексту знижує чутливість; частково це компенсується вагами ключових одиниць і консервативними порогоми, але не усувається повністю. Попри це, система забезпечує саме ті властивості, які запитують виробничі користувачі: прозорість рішень на рівні окремих лексем і фраз, відтворюваність кожного експерименту завдяки збереженню проміжних артефактів, стабільність поведінки на різномірних потоках, де тональність повинна визначатися не лише «правильно», а й однаково при повторних запусках.

Підсумовуючи, запропонований підхід консолідує інженерні та методичні засади у впорядкований конвеєр, який перетворює нестабільні короткі україномовні тексти на об'єкт керованого аналізу. Почерговість кроків: від нормалізації та допуску корпусу до виділення ключових одиниць і лексиконної агрегації забезпечує

інспектованість кожної ланки, а двоярусне тестування з каталогом помилок дозволяє не лише вимірювати якість, а й покороково її покращувати, не руйнуючи архітектурної цілісності.

### **Перелік посилань**

1. Zavalniuk I., Bohatko V., Pavlyuk O., Kukhar N. A Comparative Study of Expressive Language Usage in Contemporary Ukrainian and English Media. *Multidisciplinary Science Journal*. 2024. Vol. 6.
2. Haladzshun Z., Kunanets N., Dvorianyn P., Makarchuk O., Veretennikova N. Compound Words as a Means of Expressing Content Tonality in the Ukrainian Text Media. *Proceedings of the 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT)*. September 2021. Vol. 2. Pp. 109–114.
3. Levchuk P., Shevchuk-Kliuzheva O. Language Choice and Changes in Speech Behavior: A Study of Bilingual Ukrainians During Wartime. *Przegląd Rusycystyczny*. 2024. No. 4(188). Pp. 154–178.
4. Olenych I., Prytula M., Sinkevych O., Khamar O. System of Automatic Determination of Ukrainian Text Tone. *Proceedings of the 2021 IEEE 12th International Conference on Electronics and Information Technologies (ELIT)*. May 2021. Pp. 80–83.
5. Navalna M., Kostusiak N. M., Stanislavchuk N., Kapeliukh D., Kushch N., Demeshko I. Call Signs of Ukrainian Military in the Language of Mass Media: Functional-Semantic Aspect. *International Journal of Basic and Applied Sciences*. 2025. Vol. 14, No. 3.
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. Залуцька О.О., Молчанова М.О., Мазурець О.В., Мельник О.І., Скрипник Т.К. Метод інтелектуального аналізу емоційної тональності текстової інформації для визначення поведінкових намірів нейромережевими засобами. *Науковий журнал «Вісник Хмельницького національного університету»* серія: Технічні науки. Хмельницький, 2023. №5 (325). Т.1. С. 67-73.
8. Молчанова М.О., Мазурець О.В., Собко О.В., Віт Р.В., Назаров В.В. Алгоритм виявлення аб'юзивного вмісту в україномовному аудіоконтенті для імплементації в об'єктно-орієнтовану інформаційну систему. *Науковий журнал «Вісник Хмельницького національного університету»* серія: Технічні науки. Хмельницький, 2024. №1 (331). С. 101-106.
9. Мазурець О.В., Козенко О.В., Собко О.В. Метод автоматизованого підбору відповідей на користувацькі запитання за семантичною подібністю. *Матеріали XII Всеукраїнської науково-практичної конференції «Глушковські читання»*. Київ – 2023. с. 106-109.
10. Молчанова М.О., Мазурець О.В., Собко О.В., Кліменко В.І., Андрощук В.І. Метод нейромережевого 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.
11. 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.

12. Mazurets O., Molchanova M., Klimenko 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.
13. 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.
14. Мазурець О.В., Віт Р.В. Інтелектуальний метод виявлення цільових об'єктів предметної області за показниками семантичної зв'язності для класифікації текстової інформації. Розвитки інформаційно-керуючих систем та технологій.: монографія. Львів-Торунь : Lina-Pres, 2024. – С.223-244.
15. 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.
16. Мазурець О., Віт Р. Інтелектуальний метод виявлення цільових об'єктів предметної області для класифікації текстової інформації. Матеріали XII Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології ІУСТ-ОДЕСА-2024». 23-25.09.2024. Одеса. 2024. С.205-208.
17. 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.
18. 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.
19. 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.