MATHEMATICAL METHODS, MODELS AND INFORMATION TECHNOLOGIES IN ECONOMY

УДК 338.242
JEL C55, C63, O33, M15, M31
DOI: 10.26906/EiR.2025.3(98).3919

МЕТОДОЛОГІЧНІ ЗАСАДИ УНІВЕРСАЛЬНОГО ПІДХОДУ ДО МОВНОЇ АНАЛІТИКИ БІЗНЕС-КОМУНІКАЦІЙ НА ОСНОВІ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ

Жуковський Дмитро Миколайович*, аспірант
Український державний університет науки і технологій

**ORCID 0009-0006-1105-7434*

© Жуковський Д.М., 2025

Стаття отримана: 21.08.2025
Стаття прийнята: 16.09.2025
Стаття опублікована: 03.10.2025

Вступ. У сучасній цифровій економіці дедалі більше бізнес-моделей базуються не лише на транзакційних даних, а й на інформації, що виникає у процесі цифрових комунікацій із клієнтами [1, с.3]. Це дає підстави виокремити специфічний клас моделей юніт-економіки, у яких мовна взаємодія (діалог, повідомлення, розмова) виступає як самостійна аналітична одиниця, що підлягає кількісному та якісному виміру.

Такі моделі особливо релевантні для галузей, де процес ухвалення рішення є багатофакторним, тривалим і вимагає не лише оцінки функціональних характеристик продукту, а й високого рівня довіри, персональної консультації чи емпатії. До таких сфер належать медичні послуги, автомобільні дилерські центри, освітні проєкти з індивідуальними консультаціями, фінансові та страхові сервіси, B2B-продажі складних рішень і юридична практика. У цих випадках ключове рішення про купівлю або продовження співпраці формується після персоналізованої комунікації з представником компанії, а не лише на основі відвідування вебсайту.

В умовах цифровізації більшість взаємодій із клієнтами відбувається у форматі голосових або текстових комунікацій. Обсяг цих даних стрімко зростає, але їхня неструктурованість значно ускладнює застосування в управлінських і стратегічних процесах. Традиційні підходи до аналізу, зокрема правил-орієнтовані та словникові методи, а також класичні моделі машинного навчання, здебільшого зосереджені на окремих аспектах, такі як частотність ключових слів чи відповідність діалогу визначеним сценаріям. Це зумовлює обмеженість аналітичних висновків і не дає можливості сформувати цілісне уявлення про якість комунікацій.

Отже, виникає потреба у формуванні нової методології, здатної інтегрувати багаторівневу параметризацію аналізу та забезпечити універсальність застосування в різних секторах економіки. У цьому контексті ключову роль відіграють великі мовні моделі, які забезпечують гнучке опрацювання текстових і голосових даних у масштабах, раніше недосяжних для класичних алгоритмів.

Огляд останніх джерел досліджень і публікацій. Теоретичні основи створення систем для обробки природної мови були закладені ще в середині двадцятого століття в роботах А. Тюрінга [2, с. 433]. Праці вчених К. Чорча та Р. Мерсера заклали підгрунтя для розвитку статистичних методів у галузі обробки природної мови, які згодом стали фундаментом сучасних технологій [3, с.1]. Наукова стаття М. Бахі розкриває концепцію NLP, його еволюцію та практичне застосування у бізнес-аналітиці, підкреслюючи трансформаційний вплив на процеси прийняття рішень на основі текстових даних [4, с.3]. Групою до-

слідників у межах науково-дослідного проєкту Google Brain була опублікована стаття [5, с.6000], у якій представлено архітектуру трансформера, що стала фундаментом сучасних великих мовних моделей і нині розглядається як ключова теоретична основа їх застосування у сфері обробки природної мови.

Метою статті є обґрунтування доцільності інтеграції великих мовних моделей у методологію юніт-економіки з метою комплексного аналізу мовних комунікацій із клієнтами, визначення їх впливу на ключові бізнес-показники та розробка теоретичних засад застосування багатопараметричних систем мовної аналітики у різних секторах економіки.

Основний матеріал і результати. У сучасних умовах цифрової економіки мовна аналітика стає одним із ключових інструментів підвищення ефективності бізнес-процесів, оптимізації клієнтського досвіду та підтримки управлінських рішень. Системи мовної аналітики охоплюють широкий спектр завдань – від ідентифікації інтенцій звернень і класифікації діалогів до виявлення критичних порушень у комунікації, аналізу емоційного забарвлення висловлювань і прогнозування задоволеності клієнтів. На початкових етапах розвитку ці системи мовної аналітики будувалися переважно на правилах та словникових методах [6, с.954], що спиралися на ключові слова, регулярні вирази та вручну визначені шаблони. Попри їхню простоту, такі методи виявилися недостатньо гнучкими, оскільки погано враховували синонімію, різні мовні конструкції та потребували постійного оновлення, що ускладнювало масштабування.

Значний прогрес у розвитку мовної аналітики забезпечило впровадження методів машинного навчання, які дозволили автоматизувати обробку великих обсягів даних та підвищити точність класифікації [7, с. 332]. Найчастіше у цій сфері застосовувалися алгоритми Байєса [8, с. 387], підтримувальних векторних машин (SVM) [9, с. 122], методи найближчих сусідів [10, с. 223], ансамблеві підходи на кшталт рандомних лісів [11, с. 1063], а також нейронні мережі малої глибини [12, с. 347]. Завдяки цим підходам стало можливим враховувати приховані закономірності у даних, обробляти багатомовні набори даних текстів і частково вирішувати проблему варіативності мовних форм. Водночас їхні обмеження були пов'язані з необхідністю трудомісткого процесу створення ознак, залежністю від великих масивів навчальних прикладів та обмеженою здатністю враховувати широкий контекст діалогу, що є критично важливим для багатьох завдань мовної аналітики.

Революційною іновацією у цій сфері стало поширення великих мовних моделей, побудованих на архітектурах трансформерного типу. Завдяки багаторівневому урахуванню контексту та здатності до узагальнення LLM виявилися значно ефективнішими за традиційні методи у задачах класифікації інтенцій, аналізу тональності, виявлення прихованих смислових структур і навіть автоматичного формування рекомендацій. Їхня здатність інтегрувати галузеві знання та працювати у режимі реального часу дозволила створювати універсальні алгоритми для вирішення широкого спектра завдань мовної аналітики в охороні здоров'я, банківському секторі та автомобільному бізнесі. У порівнянні з класичними методами машинного навчання саме LLM відкривають можливість досягати високої точності без необхідності складного ручного налаштування ознак і дають змогу забезпечити масштабованість аналітичних рішень.

Автоматичне резюмування дзвінків є однією з ключових задач мовної аналітики, що спрямована на трансформацію великих обсягів неструктурованих даних у компактні та зрозумілі узагальнення [13, с. 74]. Його основне призначення полягає у стислому відтворенні змісту діалогів між клієнтами та операторами з відображенням ключових етапів комунікації: від привітання та ідентифікації клієнта до виявлення потреб, обговорення продуктів чи послуг, роботи із запереченнями та визначення підсумкових результатів взаємодії. Такий підхід дозволяє оптимізувати аналіз дзвінків і суттєво зменшити витрати часу на опрацювання розмов, зберігаючи при цьому цінність отриманої інформації.

Для підприємств автодилерського бізнесу, медичних закладів та банків комунікаційні діалоги характеризуються подібною структурою, що включає етапи привітання, виявлення потреб, презентації, опрацювання заперечень і завершення. Універсальність такої схеми забезпечує можливість застосування автоматичного резюмування як уніфікованого інструменту аналізу з подальшою адаптацією до галузевої специфіки.

Особливого значення набуває збереження у згенерованих резюме специфічних бізнес-термінів і ключових фраз, які мають вагоме значення для подальшого аналізу. У процесі роботи мовної моделі важливо, щоб назви автомобільних моделей, медичних послуг чи фінансових продуктів були відобра-

жені у резюме без змін. Саме ці маркери забезпечують можливість подальшої класифікації звернень, тематичного аналізу та формування практичних рекомендацій для підвищення ефективності бізнес-процесів.

Для реалізації цієї задачі може використовуватися універсальний промпт до LLM, який містить інструкцію: *необхідно проаналізувати текст дзвінка та скласти стисле резюме, що відображає основні етапи діалогу {етапи_діалогу} та підсумковий результат. Збережи важливі для бізнесу терміни й назви продуктів, а саме {словник термінів}, без змін*. У зазначений запит до LLM інтегруються динамічно визначені параметри для конкретного виду бізнесу, характеристика яких подається у подальшому викладі.

У змінну {етапи_діалогу} необхідно підставити етапи комунікації з урахуванням специфіки певної галузі бізнесу: для автодилерів – привітання, уточнення марки та моделі автомобіля, перевірка наявності й вартості, пропозиція тест-драйву, обговорення умов фінансування чи лізингу, завершення взаємодії; для медичних закладів – привітання, запис або узгодження прийому, опис симптомів чи скарг, уточнення результатів аналізів або діагнозу, пояснення плану лікування, завершення консультації; для банківських установ – привітання, ідентифікація потреб клієнта, презентація умов конкретного фінансового продукту, роз'яснення відсоткових ставок чи комісій, опрацювання заперечень, завершення обслуговування.

До набору параметрів {словник_термінів} належать ключові поняття та терміни, що мають зберігатися у резюмованому тексті без змін, оскільки вони формують основу змістової релевантності для бізнесу. У сфері автомобільного рітейлу такими є назви марок і моделей автомобілів, умови кредитних і лізингових програм, а також характеристики сервісних пакетів та додаткових послуг. У медичній практиці акцент робиться на коректному відтворенні назв обстежень, методів лікування та профілактичних програм, які визначають траєкторію подальшої взаємодії з пацієнтом. У банківському секторі особливого значення набуває точна фіксація фінансових продуктів, умов кредитів і депозитів, а також цифрових сервісів, що формують клієнтський досвід. Таким чином, словник термінів виступає базовим параметром, який забезпечує сталість ключових змістових елементів і уніфікацію підсумкового резюме незалежно від галузевої специфіки.

Таким чином, автоматичне резюмування дзвінків із включенням ключових термінів і фраз виступає не лише інструментом оптимізації роботи операторів та менеджерів, але й джерелом структурованих даних для глибинного аналізу. Отримані узагальнення забезпечують своєчасну адаптацію продуктового портфеля, вдосконалення маркетингових стратегій та оптимізацію сервісу.

Контроль дотримання скрипта є однією з ключових задач мовної аналітики, спрямованих на перевірку відповідності діалогу формалізованому сценарію обслуговування [14, с. 7]. Така перевірка передбачає встановлення факту проходження обов'язкових етапів, оцінювання коректності переходів між ними, використання нормативних формулювань та уникнення заборонених висловлювань, що у підсумку дозволяє сформувати висновок щодо рівня відповідності комунікаційним стандартам бізнесу. Для автоматизації цього процесу застосовується універсальна інструкція до великих мовних моделей, яка може бути подана у вигляді промпту: *проаналізуй діалог між клієнтом і оператором та перевір: чи пройдено всі обов'язкові {етапи_комунікації}, чи дотримано їх послідовність, чи використано {нормативні_формулювання}, а також чи відсутні {заборонені_висловлювання}. На основі аналізу сформулюй висновок щодо рівня відповідності діалогу корпоративним стандартам комунікації.*

Набір значень для змінної {етапи_комунікації} представлено у задачі автоматичного резюмування діалогу вище, тому, аби уникнути тавтології, у межах задачі автоматизованого контролю скрипта охарактеризуємо лише сутність параметризації для змінних {нормативні_формулювання} та {заборонені_висловлювання}. У сфері автодилерського бізнесу до нормативних формулювань належать стандартизовані вирази, спрямовані на створення довіри та підкреслення професійності, наприклад привітання клієнта, пропозиція тест-драйву чи згадка про офіційну гарантію, тоді як заборонені висловлювання охоплюють негативні оцінки продукції або некоректні порівняння з конкурентами. В охороні здоров'я нормативні формулювання закріплюють етичність і точність медичної комунікації, зокрема коректні запити про симптоми, рекомендації щодо обстежень чи пояснення плану лікування, тоді як заборонені висловлювання стосуються знецінення скарг пацієнта або заперечення серйозності захворювання. У банківській сфері нормативні фрази забезпечують прозорість і зрозумілість обслуговування, наприклад підтвердження даних клієнта, опис умов кредиту чи депозиту,

рекомендація цифрових сервісів, тоді як заборонені висловлювання вказують на відсутність знань про продукт або обмеження його доступності для клієнта.

Ідентифікація інтенції звернень є однією з базових задач мовної аналітики, що полягає у визначенні комунікативної мети клієнта під час взаємодії з оператором або автоматизованою системою [15, с. 67]. Метою цієї задачі є класифікація висловлювань не лише за їх формальною структурою, а насамперед за прагматичним змістом, тобто тим, яку дію очікує від клієнта бізнес, зокрема отримати інформацію, здійснити покупку, зафіксувати скаргу тощо. Універсальний промпт до великих мовних моделей для реалізації цієї задачі може бути сформульований так: *проаналізуй наведений діалог між клієнтом і оператором та визнач, яка з {список_можливих_інтенцій} є найбільш релевантною. Якщо серед наданого списку немає відповідної інтенції, визнач її самостійно, використовуючи зміст діалогу. Сформулюй короткий і чіткий висновок, зберігаючи нейтральний стиль викладу.*

Змінна {список_можливих_інтенцій} відображає класифікаційні категорії, які визначають мету або основну причину звернення клієнта. Її наповнення залежить від галузевої специфіки та формує основу для універсального алгоритму ідентифікації інтенцій. В автодилерському бізнесі типовими прикладами є інтенції, що стосуються уточнення моделі чи комплектації автомобіля, запису на тест-драйв, обговорення вартості, уточнення умов кредитування чи лізингу. В охороні здоров'я до основних інтенцій належать опис симптомів, прохання про консультацію, уточнення результатів обстежень, запит на призначення лікування. У банківській сфері типовими є інтенції отримати інформацію про продукт, уточнити умови договору, дізнатися про відсоткові ставки чи комісії, подати скаргу або запросити індивідуальну пропозицію. Зазначені приклади не є вичерпними й відображають лише частину можливих значень змінної {список_можливих_інтенцій}, яка може динамічно розширюватися відповідно до бізнес-контексту. Таким чином, цей параметр забезпечує як стандартизацію виявлення інтенцій, так і збереження гнучкості системи мовної аналітики у різних предметних сферах.

Тематичне тегування є однією з ключових задач мовної аналітики, спрямованою на класифікацію висловлювань і діалогів за змістовими категоріями. Мета цієї задачі полягає у формуванні системи позначень або тегів, що відображають основну тематику звернення клієнта, дозволяючи структурувати великі обсяги комунікативних даних та забезпечувати швидкий пошук релевантної інформації. На відміну від ідентифікації інтенцій, яка фокусується на комунікативній меті клієнта, тематичне тегування визначає предметну область та ключові змістові аспекти звернення.

Універсальний промпт для великих мовних моделей може бути сформульований так: *проаналізуй наведений діалог і визнач основні теми, до яких належать висловлювання клієнта. Використай {список_можливих_тем} як основу для класифікації, а якщо тема не відповідає жодному зі списку, визнач її самостійно. Подай результат у вигляді коротких тематичних тегів.*

Специфіка набору тегів для змінної {список_можливих_тем} залежить від галузі бізнесу. В автодилерському секторі типовими є теги модель автомобіля, тест-драйв, вартість, кредит/лізинг, сервісне обслуговування. В охороні здоров'я тематичні категорії охоплюють симптоми, результати аналізів, лікування, консультація лікаря, профілактика. У банківській сфері теги зазвичай відносяться до таких тем, як депозити, кредити, відсоткові ставки, комісії, онлайн-банкінг, скарги. Таким чином, змінна {список_можливих_тем} дозволяє стандартизувати класифікацію діалогів, одночасно забезпечуючи адаптацію до специфіки галузі та створюючи основу для побудови аналітичних моделей більш високого рівня.

Аналіз емоційного тону є базовою задачею мовної аналітики, що спрямована на виявлення емоційного забарвлення висловлювань клієнта й оператора у процесі діалогу. Основна мета полягає у визначенні полярності (позитивна, нейтральна чи негативна) та деталізованих станів, таких як довіра, сумнів, зацікавленість, роздратування, тривожність чи задоволення. На відміну від тематичного тегування, яке відображає зміст звернень, аналіз емоційного тону зосереджується на афективному контексті, що безпосередньо впливає на якість сервісу, рівень задоволеності та клієнтську лояльність.

Універсальний промпт для великої мовної моделі формулюється наступним чином: *проаналізуй наведений діалог і визнач емоційний тон висловлювань клієнта та оператора (позитивний, негативний, нейтральний, довіра, сумнів, зацікавленість, роздратування, тривожність, задоволення). Використай динамічні маркери ({маркери_емоцій}) для ідентифікації відповідних станів та сформулюй короткий висновок із поясненням ключових сигналів.*

Змінна {маркери_емоцій} охоплює лінгвістичні та поведінкові сигнали, які відображають специфіку галузі та слугують індикаторами для розпізнавання емоцій. У сфері автодилерського бізнесу такими

маркерами виступають згадки про інтерес до тест-драйву, сумніви щодо умов фінансування чи ентузіазм при обговоренні моделі автомобіля. В охороні здоров'я вони пов'язані з тривожністю пацієнта під час опису симптомів, недовірою до результатів діагностики або відчуттям полегшення після консультації лікаря. У банківській сфері маркери включають впевненість у вигідності умов, сумніви щодо прихованих комісій або прояви фрустрації у випадку складних процедур.

Таким чином, параметризація аналізу емоційного тону забезпечує баланс між універсальністю алгоритму та врахуванням предметної специфіки, що робить цей підхід ефективним для підвищення якості комунікацій і побудови прогнозних моделей задоволеності клієнтів у різних галузях бізнесу.

Кваліфікація лідів є прикладною задачею мовної аналітики, спрямованою на визначення рівня зацікавленості потенційного клієнта та його готовності до здійснення цільової дії. Основна мета полягає у віднесенні звернень до категорій кваліфікованих чи некваліфікованих лідів на підставі аналізу змісту діалогу, виявлення потреб, згадок про ключові параметри продукту чи послуги, а також оцінки емоційного тону. Така класифікація дозволяє автоматизувати процеси пріоритизації звернень, підвищити ефективність роботи відділів продажу й сервісу та оптимізувати витрати на залучення клієнтів.

Для реалізації такого підходу використовується універсальна інструкція до LLM-моделі: *проаналізуй діалог між клієнтом і оператором та визнач рівень готовності клієнта до угоди з урахуванням {критерії_кваліфікації}, класифікувавши його як високий, середній або низький залежно від виявлених маркерів зацікавленості й уточнювальних запитань.*

У цьому контексті змінна {критерії_кваліфікації} відображає набір умов, що характеризують готовність клієнта до наступного етапу взаємодії. Для автодилерського бізнесу такими критеріями є наявність уточнень щодо моделі автомобіля, зацікавленість у тест-драйві чи готовність обговорювати умови фінансування. В охороні здоров'я це може бути запит на консультацію, уточнення щодо конкретних обстежень або готовність розглянути запропонований план лікування. У банківській сфері критеріями кваліфікації виступають інтерес до умов кредиту чи депозиту, готовність надати персональні дані для оформлення продукту, а також запит на порівняння різних фінансових інструментів. Водночас зазначений перелік не є вичерпним: він може бути розширений у разі зміни бізнес-процесів чи появи нових ознак поведінки клієнтів, що впливають на оцінку їхнього потенціалу. Така відкритість параметризації забезпечує адаптивність моделі та її релевантність у динамічному ринковому середовищі.

Таким чином, параметризація кваліфікації лідів забезпечує гнучкість у застосуванні алгоритму в різних галузях, дозволяє стандартизувати процеси оцінювання звернень і водночас враховує предметну специфіку, що підвищує точність прогнозування комерційної цінності потенційних клієнтів.

Обробка обірваних комунікацій є однією з ключових задач мовної аналітики, що має на меті виявлення діалогів, які завершилися передчасно і не привели до досягнення цільового результату [16, с. 1824]. Основна сутність цього процесу полягає у визначенні етапу, на якому було втрачено контакт, ідентифікації причин обриву та формуванні рекомендацій щодо відновлення взаємодії або повторного звернення. У межах цього підходу розглядаються два типи чинників, що можуть спричинити перервану комунікацію. По-перше, це поведінкові причини, які пов'язані безпосередньо з реакціями клієнта, зокрема відмова від продовження розмови, зниження інтересу до продукту чи послуги, негативна реакція на пропозицію або емоційне відсторонення. По-друге, це технічні причини, які відображають функціонування інфраструктури комунікацій і визначаються не лише за змістом діалогу, але й на основі супровідних метаданих, таких як інформація про стан мережевого з'єднання, системні логи контакт-центру, записи помилок інтегрованих платформ або дані про перервані онлайн-сесії.

Для автоматизації процесу контролю застосовується універсальна інструкція до великої мовної моделі (LLM), яка формулюється наступним чином: *проаналізуй діалог клієнта й оператора разом із супровідними технічними логами та визнач, чи був він завершений коректно. Якщо комунікація перервалася передчасно, класифікуй причину як поведінкову ({поведінкові_причини}) або технічну ({технічні_причини})*. У підсумковому висновку враховуй як зміст діалогу, так і метадані, що відображають технічний стан системи. Таким чином, у межах одного алгоритму можливо здійснювати системний аналіз, що враховує як реакції клієнта, так і технічний контекст.

Практичне застосування цього підходу підтверджує його універсальність у різних галузях бізнесу. В автодилерському секторі типовими прикладами обривів є ситуації, коли діалог переривається під час обговорення вартості автомобіля, умов фінансування або після пропозиції тест-драйву. В охороні здоров'я характерними є випадки перерваних консультацій у процесі опису симптомів або після озву-

чення вартості лікування. У банківській сфері діалоги нерідко обриваються після пояснення умов кредитування або уточнення вимог до необхідних документів. Незалежно від сфери діяльності, причини обриву можна класифікувати як поведінкові або технічні, що забезпечує уніфіковану параметризацію та підвищує масштабованість аналітичних рішень. Завдяки використанню метаданих стає можливим точніше відрізняти поведінкові та технічні фактори, що дозволяє інтегрувати результати в системи управління взаємовідносинами з клієнтами й удосконалювати якість сервісу.

Можна дійти висновку, що урахування як поведінкових, так і технічних причин обірваних комунікацій дозволяє формувати комплексні моделі оцінювання ризиків та розробляти стратегії мінімізації втрат клієнтів. Це створює підґрунтя для підвищення ефективності каналів взаємодії, зменшення частки незавершених звернень та зміцнення лояльності у сфері автодилерського бізнесу, охорони здоров'я й банківського обслуговування.

Виявлення можливостей апселлінгу та кросселлінгу є прикладною задачею мовної аналітики, спрямованою на автоматизоване визначення у діалозі сигналів, які свідчать про готовність клієнта до розширення або диверсифікації покупки. Апселлінг (англ. upselling) розглядається як стратегія пропонування клієнту дорожчого, розширеного або більш функціонального продукту чи послуги, тоді як кросселлінг (англ. cross-selling) означає стимулювання придбання додаткових товарів або сервісів, що доповнюють основне замовлення. Завдяки застосуванню мовної аналітики можливо фіксувати як прямі сигнали, наприклад, уточнення щодо преміальних пакетів чи додаткових функцій, так і непрямі ознаки, зокрема порівняння варіантів, акцент на якості чи прагнення індивідуалізації. Результатом є формування рекомендацій для оператора або автоматизованої системи підтримки, що сприяє підвищенню середнього чеку, довгострокової цінності клієнта та загальної ефективності комунікацій.

Для реалізації цієї задачі застосовується універсальна інструкція до великих мовних моделей, яка може бути сформульована таким чином: *проаналізуй діалог між клієнтом і оператором та визнач, чи містить він сигнали можливості апселлінгу ({сигнали_апселлінгу}) або кросселлінгу ({сигнали_кросселлінгу}). Врахуй як прямі висловлювання клієнта, так і непрямі ознаки інтересу до додаткових чи дорожчих продуктів. Сформулюй висновок і запропонуй рекомендацію щодо подальших дій.* У цьому контексті змінна {сигнали_апселлінгу} відображає індикатори готовності клієнта обрати дорожчий або більш функціональний варіант продукту чи послуги, тоді як змінна {сигнали_кросселлінгу} фіксує ознаки зацікавленості у додаткових чи суміжних товарах і послугах, що розширюють основну покупку. Обидві змінні мають відкриту структуру і можуть гнучко параметризуватися залежно від галузевої специфіки.

Практична реалізація параметризації демонструє різні акценти у сферах бізнесу. В автодилерському секторі до сигналів апселлінгу належать згадки про дорожчу модель автомобіля чи преміальну комплектацію, тоді як сигналами кросселлінгу є інтерес до страхування, гарантійного обслуговування або аксесуарів. У сфері медичних послуг апселлінг відображається в інтересах до преміальних пакетів чи розширених програм лікування, а кросселлінг виявляється у рекомендаціях додаткових обстежень або профілактичних програм. У банківській сфері сигнали апселлінгу фіксуються через готовність перейти на преміальні тарифні пакети, тоді як сигнали кросселлінгу проявляються в інтересі до додаткових фінансових продуктів, зокрема депозитів, кредитних карток чи страхових послуг. Універсальний характер цієї задачі дозволяє розглядати апселлінг та кросселлінг як інструменти, що водночас підвищують комерційну ефективність та забезпечують стратегічну цінність клієнтської бази у будь-якій галузі бізнесу.

Аналіз переваг та недоліків продуктів у порівнянні з конкурентами є однією з ключових задач мовної аналітики, спрямованою на виявлення сильних і слабких сторін продукту або послуги шляхом їх зіставлення з альтернативними пропозиціями ринку. У межах цього процесу діалоги між клієнтами та операторами аналізуються для виокремлення згадок про характеристики як власного продукту, так і конкурентних рішень, із урахуванням контексту їх подання – позитивного чи негативного. Значну увагу приділяють частотності повторення таких згадок, емоційному забарвленню висловлювань і впливу порівнянь на готовність клієнта здійснити цільову дію, що у підсумку дозволяє формувати карту конкурентних переваг і слабких місць, придатну для інтеграції у маркетингові та продуктові стратегії. Для автоматизації цього процесу застосовується універсальний промпт до великих мовних моделей, який формулюється так: *проаналізуй діалог між клієнтом і оператором та визнач згадки про переваги {сильні_сторони} і недоліки {слабкі_сторони} продукту компанії у порівнянні з конкурентами. Склади список сильних і слабких сторін, а також вкажи фактори, що найчастіше впливали на рішення клі-*

єнтів {ключові_фактори}. Змінна {сильні_сторони} позначає характеристики продукту чи послуги, які у діалогах згадуються як конкурентні переваги, наприклад ціна, якість чи наявність додаткових сервісів. Змінна {слабкі_сторони} охоплює виявлені недоліки, що знижують привабливість продукту, зокрема високі витрати, обмежений функціонал або недостатній рівень сервісу. Змінна {ключові_фактори} відображає ті аспекти, які безпосередньо впливають на прийняття клієнтських рішень, включно з умовами фінансування, швидкістю обслуговування чи якістю післяпродажної підтримки. Усі ці змінні мають гнучкий характер і можуть адаптуватися до специфіки конкретної галузі бізнесу. Так, у сфері автомобільного дилерства аналітика зосереджується на вартості, якості сервісного обслуговування, доступності кредитних програм та наявності гарантійних умов; у медичній сфері головними параметрами є рівень кваліфікації лікарів, технологічне оснащення, якість післялікарняного супроводу та цінова політика порівняно з іншими закладами; у банківському секторі провідне значення мають умови кредитування, депозитні ставки, швидкість обслуговування та доступність цифрових каналів. Таким чином, параметризація універсального промпту до LLM дає змогу перетворювати дані з комунікацій у цінні практичні інсайти, що забезпечують удосконалення продуктового портфеля, оптимізацію цінової політики та посилення конкурентних позицій компанії, а сам процес аналізу постає як універсальний і водночас гнучкий інструмент підтримки стратегічних управлінських рішень.

Висновки. У дослідженні обґрунтовано доцільність використання великих мовних моделей (LLM) для вирішення широкого спектра задач мовної аналітики, що мають критичне значення для бізнесу. Показано, що традиційні rule-based методи та класичні алгоритми машинного навчання не забезпечують належної гнучкості та адаптивності, тоді як LLM відкривають можливості для побудови багатопараметричних алгоритмів, здатних інтегруватися у різні галузі бізнесу з урахуванням їхньої специфіки. Запропонований підхід базується на уніфікованій структурі інструкцій до LLM із подальшою параметризацією, що дозволяє стандартизувати аналітичні процедури й водночас адаптувати їх до особливостей конкретного бізнес-контексту.

Практична цінність роботи полягає у формуванні єдиного методологічного каркасу для аналізу комунікацій у ключових сферах бізнесу, зокрема автомобільному секторі, медицині та банківських послугах. Інтеграція таких моделей у межах CRM-систем, модулів контролю якості, управлінських панелей та звітності для регуляторів забезпечує оптимізацію використання ресурсів, підвищення якості сервісу та зростання ефективності юніт-економіки підприємств. Використання динамічних параметрів робить системи мовної аналітики універсальними, а також створює підґрунтя для масштабування запропонованого підходу у суміжних сферах бізнесу, де комунікаційні процеси є критично важливими, що підтверджує загальну гнучкість і довгострокову практичну значущість розробленої концепції.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Calderon-Monge E. & Ribeiro-Soriano D. The role of digitalization in business and management: a systematic literature review. *Review of Managerial Science.* 2023. P. 1–43.
2. Turing A. M. Computing Machinery and Intelligence. *Mind.* 1950. Vol. 59. № 236. P. 433–460.
3. Church K. and Mercer R. Introduction to the special issue on computational linguistics using large corpora. *Computational Linguistics*. 19 (1). 1993. P. 1–24.
4. Bahja M. Natural Language Processing Applications in Business. *E-Business – Higher Education and Intelligence Applications*. 2020. P. 1–25.
5. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł. and Polosukhin I. (2017) Attention is All You Need.*Advances in Neural Information Processing Systems*. № 30. P. 5998–6008.
6. Kang N., Singh B., Afzal Z. Using rule-based natural language processing to improve disease normalization in biomedical text. *Journal of the American Medical Informatics Association*. № 20 (5). 2013. P. 953–958.
7. Smith J., Brown P., Taylor R. Rule-based natural language processing for clinical practice. *Journal of Clinical Informatics*. № 38 (4). 2024. P. 332–337.
8. Finkel J.R., Grenager T., Manning C. Bayesian analysis in natural language processing. *Computational Linguistics.* № 44 (1). 2018. P. 187–218.
9. Hou Y., Huang J. Natural language processing for social science research: A comprehensive review. Chinese Journal of Sociology. № 11(1). P. 121–157.
10. Ge, Y., Al-Garadi, M.A. and Sarker, A. Data Augmentation with Nearest Neighbor Classifier for Few-Shot Named Entity Recognition. *Studies in Health Technology and Informatics*. № 314. 2024. P. 220–225.
11. Wang J., Li Y., Chen Q. A semantics aware random forest for text classification. *Pattern Recognition Letters.* № 126. 2019. P. 1061–1070.
12. Goldberg Y. A primer on neural network models for natural language processing. *Journal of Artificial Intelligence Research*. № 60. 2017 .P. 345–420.

13. Jayatilleke N., Weerasinghe R., Senanayake N. Advancements in Natural Language Processing for Automatic Text Summarization. *International Conference on Computer Systems (ICCS).* 2024. P. 74–84.

14. Barr V., Klavans J.L. Verification and Validation of Language Processing Systems. *Proceedings of the First Meeting of the North American Chapter of the Association for Computational Linguistics*. 2001. P. 6–14.

15. Saranraj R., Vasanthakumar A., Arivoli T. Pretrained Natural Language Processing Model for Intent Recognition. *Atlantis Press.* P. 66–76.

16. Pan Z., Bai K., Wang Y., Zhou L., Liu X. Improving Open-Domain Dialogue Systems via Multi-Turn Incomplete Utterance Restoration. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2019 P. 1824–1833.

УДК 338.242
JEL C55, C63, O33, M15, M31

Жуковський Дмитро Миколайович, аспірант, Український державний університет науки і технологій.
Методологічні засади універсального підходу до мовної аналітики бізнес-комунікацій на основі великих мовних моделей.

У статті обґрунтовується доцільність застосування великих мовних моделей LLM (англ. Large Language Model) як принципово нового інструменту в задачах мовної аналітики для бізнесу. У роботі систематизовано прикладні задачі мовної аналітики, що реалізуються з використанням LLM: автоматичне резюмування діалогів, контроль дотримання скрипта, ідентифікація інтенцій звернень, тематичне тегування, аналіз емоційного тону, кваліфікація лідів, виявлення можливостей апселлінгу та кросселлінгу, обробка обірваних комунікацій та аналіз задоволеності клієнтів. Для кожної задачі запропоновано універсальні промпти із динамічними параметрами, що дозволяє адаптувати інструкції до специфіки окремих галузей бізнесу. Особливий акцент зроблено на трьох секторах із критичною залежністю від мовної аналітики – автодилерському, медичному та банківському.

Ключові слова: цифрова економіка, системний аналіз, штучний інтелект, обробка природньої мови, мовна аналітика, юніт економіка, великі мовні моделі, управлінський облік, методи машинного навчання.

UDC 338.242
JEL C55, C63, O33, M15, M31

Dmytro Zhukovskyi, Postgraduate Student, Ukrainian State University of Science and Technology.
Methodological foundations of a universal approach to business communication language analytics based on large language models.

The article substantiates the feasibility of using large language models (LLMs) as a fundamentally new tool in business-oriented language analytics tasks. Unlike traditional rule-based approaches and dictionary methods, which are characterized by low flexibility and limited ability for contextual interpretation, as well as classical machine learning algorithms that require large volumes of labeled data and quickly lose relevance in a dynamic business environment, LLMs provide comprehensive multiparametric analysis of communications. Their scientific novelty lies in the ability to integrate semantic, contextual, and emotional dimensions of interaction, while also taking into account industry-specific features without the need to develop narrowly specialized models. The study systematizes applied tasks of language analytics implemented through LLMs, including automatic dialogue summarization, script compliance control, intent identification, thematic tagging, sentiment analysis, lead qualification, upselling and cross-selling opportunity detection, handling of interrupted communications, and customer satisfaction analysis. For each task, universal prompts with dynamic parameters are proposed, enabling adaptation of instructions to the specifics of individual business domains. Particular emphasis is placed on three sectors with critical dependence on language analytics – automotive dealerships, healthcare institutions, and banking services. This approach makes it possible to combine the standardization of analytical procedures with domain-specific detailing for each industry. The practical value of the study lies in the development of a unified methodological framework for integrating LLM-based language analytics into CRM systems, quality control modules, management dashboards, and regulatory reporting. The use of dynamic parameters ensures both universality and scalability of the approach, opening prospects for its application in a wider range of business areas where client communication is a key factor in ensuring unit economics efficiency and enhancing enterprise competitiveness.

Key words: digital economy, systems analysis, artificial intelligence, natural language processing, language analytics, unit economics, large language models, managerial accounting, machine learning methods.