

РАЛІК ІГОР

Тернопільський національний технічний університет
<https://orcid.org/0009-0007-9769-3331>
e-mail: ihortntu@gmail.com

ГЕНЕРАТИВНІ МОДЕЛІ ДЛЯ СТВОРЕННЯ РЕЗОНАНСНИХ СЦЕНАРІЇВ КОМУНІКАЦІЇ У CRM-СИСТЕМАХ

У статті представлено модульну архітектуру CRM, яка поєднує емоційно-когнітивний аналіз клієнта з динамічним формуванням персоналізованих сценаріїв комунікації за допомогою великих мовних моделей. Ключовим елементом є механізм структурованого пром프트у, який включає рольові інструкції, історію діалогу, опис стану клієнта та бізнес-правила. Метою дослідження є розроблення й експериментальна оцінка підходу до генерації сценаріїв, що одночасно враховують емоційні та когнітивні параметри користувача для підвищення резонансності відповідей і якості взаємодії.

Стан клієнта подано у вигляді двох нормованих векторів – емоційного та когнітивного, що дозволяє компактно описати його поточні реакції й наміри. На основі цих векторів формується динамічний структурований промпт, який використовується для генерації релевантних альтернативних сценаріїв відповіді. Кожен сценарій оцінюється через резонансну функцію, а оптимальний варіант обирається селектором стратегій з урахуванням резонансу та попередньої ефективності. Баланс між емоційним і когнітивним компонентами динамічно коригується RL-агентом.

Робота формалізує двокомпонентне (емоційне та когнітивне) представлення стану клієнта для генеративної персоналізації, вводить резонансну функцію як метрику відповідності сценарію та вперше поєднує динамічне формування пром프트ів з емоційно-когнітивним аналізом клієнта для побудови високорезонансних сценаріїв комунікації.

Ефективність системи оцінювалась у симульованому середовищі на основі 500 діалогових сесій із варіативними емоційно-когнітивними профілями клієнтів. Порівняння з базовою (baseline) моделлю продемонструвало істотне покращення ключових метрик: середній показник резонансу зріс з 0.62 до 0.71, частка високорезонансних сценаріїв збільшилася з 28% до 43%, а кількість сценаріїв з низьким резонансом зменшилася удвічі – з 22% до 11%. Отримані результати підтверджують, що динамічно сформовані генеративні сценарії забезпечують вищу персоналізацію, стабільність поведінки та потенційне зростання конверсії.

Запропонована модель демонструє ефективність генеративних алгоритмів у CRM-системах, здатних адаптувати стиль і зміст комунікації до поточного емоційно-когнітивного стану клієнта в реальному часі. Механізм динамічного пром프트ування та резонансного вибору сценаріїв може бути інтегрований у чат-боти, контакт-центри та сервісні платформи, підвищуючи якість автоматизованих відповідей і зменшуючи частку нерелевантних повідомлень без змін у бізнес-логіці.

Ключові слова: CRM-система, генеративна модель, велика мовна модель, емоційно-когнітивний резонанс, когнітивний аналіз, емоційний аналіз.

RALIK IHOR

Ternopil National Technical University

GENERATIVE MODELS FOR CREATING RESONANT COMMUNICATION SCENARIOS IN CRM SYSTEMS

The article presents a modular CRM architecture that combines emotional-cognitive client analysis with dynamic generation of personalized communication scenarios using large language models. A key element is the structured prompt mechanism, which includes role instructions, dialogue history, a description of the client's state, and business rules. The aim of the study is to develop and experimentally evaluate an approach to scenario generation that simultaneously accounts for the user's emotional and cognitive parameters to enhance the resonance of responses and the quality of interaction.

The client's state is represented as two normalized vectors (emotional and cognitive) which allow for a compact description of current reactions and intentions. Based on these vectors, a dynamic structured prompt is formed and used to generate relevant alternative response scenarios. Each scenario is evaluated through a resonance function, and the optimal option is selected by a strategy selector that considers resonance and prior effectiveness. The balance between emotional and cognitive components is dynamically adjusted by an RL agent.

The work formalizes a two-component (emotional and cognitive) representation of the client's state for generative personalization, introduces the resonance function as a metric of scenario relevance, and for the first time combines dynamic prompt formation with emotional-cognitive client analysis to build high-resonance communication scenarios.

System effectiveness was evaluated in a simulated environment based on 500 dialogue sessions with variable emotional-cognitive client profiles. Comparison with a baseline model demonstrated significant improvements in key metrics: the average resonance score increased from 0.62 to 0.71, the share of high-resonance scenarios rose from 28% to 43%, and the number of low-resonance scenarios was halved – from 22% to 11%. The results confirm that dynamically generated scenarios provide higher personalization, behavioral stability, and potential conversion growth.

The proposed model demonstrates the effectiveness of generative algorithms in CRM systems capable of adapting the style and content of communication to the client's current emotional-cognitive state in real time. The mechanism of dynamic prompting and resonance-based scenario selection can be integrated into chatbots, contact centers, and service platforms, improving the quality of automated responses and reducing the share of irrelevant messages without changes to business logic.

Keywords: CRM system, generative model, large language model, emotional-cognitive resonance, cognitive analysis, emotional analysis.

Стаття надійшла до редакції / Received 13.12.2025
Прийнята до друку / Accepted 11.01.2026
Опубліковано / Published 29.01.2026



This is an Open Access article distributed under the terms of the [Creative Commons CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

© Ралік Ігор

Постановка проблеми

Сучасні CRM-системи демонструють високий рівень автоматизації комунікацій, однак здебільшого залишаються обмеженими у здатності враховувати психологічні характеристики клієнта під час взаємодії.

Переважаючі рішення оперують переважно статичними сценаріями або поверхневими моделями емоцій, що не дозволяє досягти глибокої персоналізації та справжнього резонансу між повідомленням і поточним станом користувача. Враховуючи існуючий науково-практичний розрив, виникає потреба у створенні підходу, який забезпечить комплексне врахування емоційних та когнітивних чинників клієнта, а також здатний адаптивно формувати комунікаційні сценарії засобами генеративних мовних моделей.

Саме тому є необхідною розробка та експериментальна оцінка генеративної моделі створення сценаріїв взаємодії у CRM, яка враховує емоційно-когнітивний стан клієнта ($E_c + C_c$), забезпечує формування множини можливих відповідей та дозволяє обирати оптимальний сценарій за критерієм резонансності. Для досягнення поставленої мети передбачається вирішення таких ключових завдань:

- формалізувати модель подання стану клієнта у вигляді двох нормованих векторів – емоційного та когнітивного компонентів;
- описати механізм побудови динамічного промпта для LLM із використанням контексту, історії взаємодій та бізнес-обмежень;
- реалізувати механізм генерації множини сценаріїв та визначення резонансного показника R_s для їх ранжування;
- розробити архітектуру селекції оптимальної відповіді з урахуванням показників якості та історичної ефективності;
- провести експериментальне порівняння запропонованого підходу з базовим (шаблонним) сценарієм і визначити його вплив на показники взаємодії та конверсії.

Таким чином, центральним дослідницьким завданням роботи є побудова цілісної моделі генеративної комунікації у CRM, яка адаптивно налаштовує стиль, структуру та інформаційне навантаження відповіді системи відповідно до зміни емоційного та когнітивного профілю користувача. Досягнення сформульованих цілей дозволить підвищити якість автоматизованого діалогу, скоротити частку нерелевантних відповідей та створити передумови для впровадження резонансно-орієнтованих CRM-систем нового покоління.

Аналіз останніх досліджень та публікацій

Наукові дослідження останніх років демонструють стійку тенденцію до інтеграції систем управління взаємовідносинами з клієнтами (CRM) із технологіями штучного інтелекту, зокрема моделями глибинного навчання[1] та великими мовними моделями (LLM)[2]. У класичних підходах CRM-системи орієнтовані переважно на сегментацію клієнтів, автоматизацію бізнес-процесів та описово-аналітичні функції, тоді як емоційно-когнітивний стан користувача здебільшого ігнорується або враховується поверхово. Більшість промислових реалізацій використовують механізми, які визначають тональність висловлювань, проте не здатні відстежувати когнітивні наміри, рівень впевненості чи інформаційні потреби клієнта у динаміці діалогу.

Публікації, присвячені застосуванню LLM для персоналізації комунікацій, здебільшого фокусуються на генерації маркетингових повідомлень, оптимізації відповіді чат-ботів або підвищенні UX-взаємодії шляхом контекстної адаптації діалогу[3]. Однак у більшості робіт відсутня формалізована модель представлення користувача, яка б одночасно враховувала два ключові параметричні компоненти – емоційний (E_c) та когнітивний (C_c) стан. Наявні підходи переважно зводяться до класифікації емоцій за валентністю та інтенсивністю[4], тоді як когнітивні параметри (information need, goal orientation, reasoning load) розглядаються окремо або зовсім не моделюються.

В останніх публікаціях відзначається зростання інтересу до моделей адаптивної комунікації, заснованих на рекурентному навчанні (RL) та багатоканальних векторних репрезентаціях клієнта. Проте більшість робіт торкаються лише окремих аспектів такої системи, або RL для оптимізації сценаріїв, або LLM як генератора реакцій[5], що не створює цілісного, резонансно-орієнтованого механізму. Саме тому запропонована в статті двовекторна модель клієнтського стану ($E_c + C_c$) та резонансна метрика R_s виступають концептуально новим рішенням у порівнянні з існуючими дослідженнями.

Важливо відзначити, що сучасні CRM рідко використовують історичні діалогові дані не лише як текстовий контекст, але і як параметр для оптимізації стратегії діалогу. Більшість комерційних систем обмежуються методами ранжування сценаріїв заснованої на правилах, у той час як LLM у поєднанні з RL дають можливість не просто відтворювати попередні реакції[6], а формувати динамічно керований генеративний сценарій, що здатний змінювати стиль, тон та інформаційний рівень відповідно до зміни емоційно-когнітивного профілю користувача.

Таким чином, проведений огляд наукових праць показує, що хоча у сфері CRM активно розвиваються методи персоналізації, питання комплексного врахування емоційних і когнітивних характеристик залишається недостатньо висвітленим. Дана стаття заповнює цей науковий розрив, пропонуючи формалізацію резонансної комунікаційної моделі та демонструючи її ефективність порівняно з шаблонними сценаріями. Саме інтеграція LLM, когнітивних індикаторів, векторних емоційних станів та RL-оптимізації формує новий напрям подальших досліджень і відкриває перспективи розвитку інтелектуальних CRM систем наступного покоління.

Мета та завдання дослідження

Розробити та експериментально оцінити генеративний модуль, який формує комунікаційні сценарії, адаптовані до емоційних та когнітивних параметрів клієнта та забезпечить найбільший емоційно-когнітивний резонанс, перевірити ефективність запропонованої моделі порівняно з шаблонними сценаріями.

Виклад основного матеріалу

Запропонована CRM-система реалізує модульний підхід, що дозволяє розділити функціональні блоки,

забезпечити масштабованість та інтеграцію різних типів алгоритмів – від моделей глибинної обробки мови до оптимізаційних стратегій навчання з підкріпленням. Робота системи організована у вигляді послідовного ланцюга модулів, кожен з яких виконує чітко визначені завдання:

- Модуль збору даних – відповідає за приймання та структурування вхідних повідомлень від клієнта.
- Модуль емоційно-когнітивного аналізу клієнта - генерує вектори емоційного E_c та когнітивного C_c станів клієнта, які формують математичне представлення його поточного психологічного стану.
- Модуль генерації персоналізованих сценаріїв - на основі векторів психологічного стану клієнта і діалогового контексту, використовуючи генеративні моделі створює набір альтернативних сценаріїв відповіді $S = \{S_1, S_2, \dots, S_n\}$.
- ESMR-аналітик – обчислює показники резонансу між станом клієнта та кожним зі сценаріїв.
- Резонансний селектор стратегій – обирає оптимальний сценарій взаємодії з урахуванням резонансу, історичних даних про ефективність сценаріїв та контексту попереднього діалогу
- RL-агент – коригує вагові коефіцієнти, що визначають баланс між емоційним та когнітивним компонентами так, щоб максимізувати ефективність вибраних сценаріїв у подальших кроках.

У системі адаптивної резонансної персоналізації ключовим елементом є формалізація психологічного стану клієнта у вигляді двох векторів: емоційного та когнітивного. Таке подання забезпечує можливість математичного аналізу, порівняння та використання стану клієнта при генерації сценаріїв комунікації.

Емоційний вектор клієнта (E_c) відображає поточний афективний стан клієнта та ґрунтується на оцінці тексту повідомлення за допомогою NLP-моделей емоційного аналізу. Всі показники нормуються у межах $[0; 1]$, де 0 означає відсутність ознаки, а 1 – її максимальну інтенсивність. До складу вектора входять такі компоненти:

- Valence (позитивність/негативність) – загальний емоційний тон висловлювання.
- Arousal (рівень емоційного збудження) – спокійне/напружене або нейтральне становище.
- Trust (довіра) – оцінка рівня готовності клієнта сприймати пропозиції системи.
- Engagement (залученість) – інтенсивність прояву інтересу до теми.
- Frustration (роздратування/незадоволеність) – маркер негативної реакції на сервіс чи контент.
- Uncertainty (емоційна невпевненість) – ознака емоційної нестабільності або вагання.

Ці параметри в сукупності визначають афективний профіль клієнта, який впливає на вибір емоційного стилю сценарію відповіді.

Когнітивний вектор клієнта (C_c) формує уявлення про раціональний, інформаційний та мотиваційний стан клієнта. Він описує не емоції, а мисленнєві процеси, пов'язані з прийняттям рішень. Компоненти когнітивного вектора також нормуються в діапазоні $[0; 1]$, а його структура включає:

- Decision Readiness (готовність до рішення) – наскільки клієнт близький до виконання цільової дії.
- Information Need (потреба в додатковій інформації) – бажання уточнити, отримати приклади, деталі або порівняння.
- Confidence Level (рівень визначеності) – ступінь упевненості у власному рішенні чи погляді.
- Goal Orientation (спрямованість на ціль) – наскільки запит клієнта має практичну мету.
- Comparative Thinking (схильність до порівняння) – орієнтація на аналіз варіантів, переваг та недоліків.
- Cognitive Load (когнітивне навантаження) – рівень складності, який готовий сприймати клієнт (просте/детальне пояснення).

Такий підхід дозволяє моделі оцінювати не лише настрій користувача, а й його раціональні наміри, що підвищує точність персоналізації.

Оцінювання відповідності сценарію здійснюється за резонансною функцією: $R_s = \alpha \cdot \text{sim}(E_c, E_s) + \beta \cdot \text{sim}(C_c, C_s)$, де:

- E_c – вектор, що описує поточний емоційний стан клієнта;
- E_s – емоційний вектор відповіді системи;
- C_c – вектор, що описує поточний когнітивний стан клієнта;
- C_s – когнітивний вектор відповіді системи;
- $\text{sim}(x, y)$ – косинусна функція подібності між векторами;
- α, β – адаптивні вагові коефіцієнти, які визначають відносну значущість емоційної та когнітивної відповідності.

Генерація персоналізованих сценаріїв відповіді у запропонованій CRM-системі здійснюється за допомогою LLM-модуля, робота якого визначається правильно сформованим промптом. У даній моделі промпт не є статичною текстовою інструкцією, а формується динамічно, виходячи зі стану клієнта, оціненого векторами E_c та C_c , а також параметрами моделі резонансу. Модуль персоналізації створює структурований промпт P_t , який передається генеративній мовній моделі та містить усю необхідну інформацію для побудови релевантних та контекстно обґрунтованих відповідей. Промпт формується як композиція кількох інформаційних блоків: $P_t = P_{\text{context}} + P_{\text{history}} + P_{\text{state}} + P_{\text{constraints}} + P_{\text{business}}$, де:

- P_{context} – рольова інструкція, що визначає функцію генеративної моделі у межах CRM-системи;
- P_{history} – релевантний фрагмент попередньої взаємодії (останні повідомлення клієнта);
- P_{state} – опис емоційного та когнітивного стану клієнта у текстовій формі;
- $P_{\text{constraints}}$ – вимоги до стилю, кількості сценаріїв, заборони та формальні обмеження;

• $P_{business}$ – доменні знання компанії, які завантажуються до системного промпту одноразово й надалі доступні моделі у кожній ітерації генерації.

Ця структура забезпечує одночасну контекстну інформованість, емоційно-когнітивну чутливість та відповідність бізнес-правилам.

Блок $P_{context}$ визначає базову роль та функціональну поведінку генеративної моделі в межах CRM-системи. Він встановлює, що модель виступає як спеціалізований модуль формування сценаріїв відповіді, основним завданням якого є генерування текстів, узгоджених з емоційним та когнітивним станом клієнта. У цьому блоці задаються загальні принципи роботи системи, включно з орієнтацією на персоналізацію, дотриманням комунікативних норм та уніфікованим стилем взаємодії. Також $P_{context}$ забезпечує сталість поведінки генеративної моделі між різними сесіями, оскільки фіксує сталі вимоги: адаптивність, відповідність контексту діалогу, уникнення суперечливих тверджень, підтримання коректного тону.

Історичний блок $P_{history}$ містить попередні повідомлення клієнта, які формують контекст діалогу. Використання лише короткого фрагмента історії забезпечує базову контекстну чутливість, однак обмежує здатність моделі враховувати логіку розвитку взаємодії та зміни у стані клієнта. З огляду на те, що сучасні генеративні моделі здатні опрацьовувати великі контекстні вікна[7-8], доцільним є включення повної історії поточного діалогу, що забезпечує глибшу когнітивну узгодженість та підвищує якість сформованих сценаріїв відповіді. Повна історія діалогу має такі переваги:

• Глибша контекстуальна релевантність – модель бачить весь хід обговорення та уникає суперечностей.

• Коректне відстеження змін емоційного та когнітивного стану – коливання параметрів стають інтерпретованими у часовому контексті.

• Зменшення ризику поверхневих відповідей – модель може враховувати попередні уточнення, очікування, стиль спілкування клієнта.

• Вища якість сценаріїв – особливо важливо для B2B, довготривалих консультацій або ситуацій, де потрібна складна аргументація.

Для оптимізації було використано підхід із плаваючим вікном: $P_{history} = \{ \text{усі повідомлення, якщо } L \leq L_{max}; \text{останні } k \text{ повідомлень, якщо } L > L_{max} \}$, де:

• L – загальна довжина діалогу.

• L_{max} – максимальна довжина контексту моделі.

• k – межа компромісу між контекстністю та продуктивністю.

Отже, система передає в генеративну модель усю історію діалогу в межах доступного контекстного вікна, а обмеження застосовуються лише для дуже довгих розмов. Це суттєво підвищує якість сценаріїв та релевантність рекомендацій, особливо в CRM-контекстах, де важливо враховувати не окремі репліки, а поточний стан діалогу як цілісного процесу.

Поточний стан клієнта подається у вигляді перетворених параметрів емоційного та когнітивного векторів: $P_{state} = gemo(Ec) + gcog(Cc)$, де:

• $gemo()$ – функція текстового перетворення емоційного вектора,

• $gcog()$ – функція текстового перетворення когнітивного вектора.

Для коректної роботи генеративної моделі дані функції перетворюють вектори у природномовний опис: “Емоційний стан: valence = 0.72; arousal = 0.41; trust = 0.84; engagement = 0.88; frustration = 0.10; uncertainty = 0.22. Когнітивний стан клієнта: decision readiness = 0.63; information need = 0.52; confidence = 0.61; goal orientation = 0.74; comparative thinking = 0.48; cognitive load = 0.42.”. Отримані описи включаються у блок P_{state} та визначають стилістичні й змістові характеристики майбутніх сценаріїв.

Блок $P_{constraints}$ визначає набір загальних правил і методологічних інструкцій, які регулюють поведінку генеративної моделі під час формування набору сценаріїв відповіді. Цей блок не залежить від конкретного бізнес-середовища і відображає універсальні вимоги до формату, стилю, безпеки та релевантності контенту. Він забезпечує контрольованість генеративного процесу та запобігає появі небажаних або некоректних відповідей. У межах модульної генеративної архітектури CRM цей блок відіграє роль метарегулятора, який визначає рамки, в яких модель інтерпретує вхідні дані, та стандарти, яких вона повинна дотримуватися під час побудови сценаріїв. Концептуально блок складається з чотирьох основних підкомпонентів: стилістичних інструкцій, обмежень безпеки, форматних вимог та змістових правил, пов’язаних із релевантністю відповіді.

Стилістична частина блоку $P_{constraints}$ визначає загальний тон, мовну манеру та допустимий рівень деталізації відповідей. Її мета забезпечити однорідність комунікації, незалежно від того, який сценарій був згенерований. До стилістичних вимог входять: використання ввічливої, нейтральної та ненав’язливої мови; уникання надмірно технічної термінології, якщо когнітивний стан клієнта свідчить про низький рівень готовності до складної інформації; рекомендація формувати відповіді у структурі “коротке пояснення – основне повідомлення – уточнювальне запитання”; уникнення емоційних перебільшень або надмірної експресивності. Стилістичні інструкції забезпечують плавність взаємодії та сприяють формуванню передбачуваної поведінки моделі в різних діалогових ситуаціях.

Друга складова блоку $P_{constraints}$ – це правила, спрямовані на забезпечення достовірності та коректності генерованих відповідей. У CRM-системах ці вимоги мають особливо велике значення, оскільки модель взаємодіє з клієнтом у реальному часі та не повинна створювати помилкових очікувань або вводити клієнта в

оману[9]. До цього підкомпонента належать: заборона на надання неправдивих гарантій або успіхів (“гарантовано отримаєте завтра”, “ми точно зможемо вирішити ваше питання”); обмеження на інтерпретацію юридичної чи договірної інформації; вимога уникати припущень щодо намірів, особистих даних або фінансового становища клієнта; обмеження щодо використання чутливої інформації; необхідність формувати тільки перевірені та нейтральні твердження. Усі ці обмеження спрямовані на мінімізацію ризиків, пов’язаних з автоматизованою генерацією контенту.

Третя частина блоку $P_{constraints}$ встановлює структуру вихідних сценаріїв, що генеруються моделлю. У більшості випадків система вимагає побудови строго певної кількості сценаріїв, кожен з яких має відповідати стандартизованій формі. Форматні вимоги можуть передбачати: фіксовану кількість сценаріїв; внутрішню структуру кожного сценарію (вступний контекст, основна відповідь, уточнювальне запитання); дотримання лаконічності; заборону відхилятися від визначеного шаблону. Ці вимоги мають на меті забезпечити однорідність сценаріїв, їх структурованість та прогнозованість.

Четверта частина блоку регулює те, яким чином сценарії повинні узгоджуватися з оціненим емоційним та когнітивним станом клієнта. Обмеження не задають конкретного тексту, але визначають принципи: у разі високого рівня Frustration сценарії мають акцентувати увагу на зниженні напруги, вибаченнях, уточненні проблеми; за високого Information Need сценарії мають містити детальніші пояснення або приклади; при низькому Decision Readiness не допускається надмірний тиск або прямі заклики до дії; за високого Engagement дозволяється більш активна пропозиційна частина.

Блок $P_{business}$ визначає сукупність правил, обмежень, знань та інструкцій, специфічних саме для конкретної організації або домену, в якому розгорнута CRM-система. На відміну від універсального блоку системних обмежень, блок бізнес-правил є доменозалежним компонентом, який може істотно змінюватися залежно від: типу компанії (банківська сфера, e-commerce, онлайн-освіта); видів послуг або товарів; політик, регламентів та внутрішніх процедур; бізнес-цілей (утримання, продаж, технічна підтримка).

Блок $P_{business}$ містить інформацію, яку модель повинна враховувати незалежно від конкретного діалогу, а саме:

- опис актуальних продуктів, тарифів, послуг;
- правила повернення товарів;
- гарантійні умови;
- політику підтримки клієнтів;
- допустимі формулювання пропозицій;
- юридичні обмеження;
- регламентовані часові рамки (доставка, обробка замовлень).

Усі ці знання не змінюються від діалогу до діалогу, тому завантажуються одноразово як системний промпт, що застосовується до всіх подальших викликів моделі. Це дає можливість: уникнути дублювання великих об’ємів тексту в кожному запиті; зменшити витрати на токени; забезпечити стабільну та контрольовану поведінку моделі; гарантувати повну відповідність бізнес-політиці[10].

Таким чином, блок $P_{business}$ створює фіксовану основу знань, з якою всі згенеровані сценарії повинні бути сумісні. Під час генерації сценаріїв він визначає межі, в яких модель може пропонувати рішення. Зокрема, він забезпечує: відповідність кожного сценарію актуальним характеристикам продуктів; коректність формулювань, які клієнт може інтерпретувати як комерційну пропозицію; відсутність тверджень, що суперечать регламентам компанії; узгодженість стилю взаємодії з клієнтом у межах бренду. Завдяки цьому система забезпечує функціональну, юридичну та операційну надійність у процесі взаємодії з клієнтом.

Для оцінювання ефективності розробленого модуля генерації сценаріїв, адаптованих до емоційно-когнітивного стану клієнта, було проведено низку експериментів у рамках побудованого прототипу CRM-системи. Основною метою експериментів була перевірка здатності генеративної моделі створювати сценарії, які узгоджуються з емоційним та когнітивним профілями клієнта; впливу динамічно сформованих промптів на резонансність сценаріїв; переваги генеративних сценаріїв над шаблонними з точки зору резонансу.

Експерименти були проведені у середовищі симуляції, де клієнтські повідомлення, їх емоційні й когнітивні характеристики та подальші реакції формувалися спеціалізованою мовною моделлю. Такий підхід забезпечив контрольованість умов, відтворюваність сценаріїв та можливість масштабного тестування без залучення реальних користувачів.

Система працювала у повному циклі, що включає всі модулі архітектури:

- аналіз вхідного повідомлення,
- формування емоційного та когнітивного векторів,
- динамічне формування промпту P_i ,
- генерацію набору сценаріїв,
- обчислення резонансу ESMR-аналітиком,
- вибір сценарію селектором стратегій,
- моделювання реакції клієнта генеративною моделлю,
- повторне оновлення стану та запуск нового циклу.

Таким чином було перевірено не окремий модуль, а узгоджену роботу всієї системи.

Параметри симуляції:

- Кількість сесій: 500.
- Середня довжина діалогу: 8 реплік.
- Кількість сценаріїв, що генерує LLM на кожному кроці: 5.
- Обсяг контексту: повна історія діалогу в межах контекстного вікна моделі.
- Baseline-модель: статичні сценарії, не прив'язані до Ес та Сс.
- Експериментальна модель: генеративні сценарії на основі динамічних промптів.

Для кожної сесії фіксувалися такі показники:

- резонанс R кожного сценарію;
- середній резонанс $\bar{R}$ для набору сценаріїв;
- частка високоякісних сценаріїв ($R > 0.75$);
- частка низькорезонансних сценаріїв ($R < 0.4$);
- зміна емоційних та когнітивних параметрів після застосування сценарію.

У ході експериментів було здійснено порівняння роботи базової CRM-моделі, яка використовує шаблонні сценарії, і запропонованої генеративної моделі, що формує відповіді на основі структурованого промпту та враховує емоційно-когнітивний стан клієнта. На рисунку 1 показано розподіл резонансу для базової та генеративної моделей, а в таблиці 1 наведено ключові показники, за якими проводилось оцінювання.

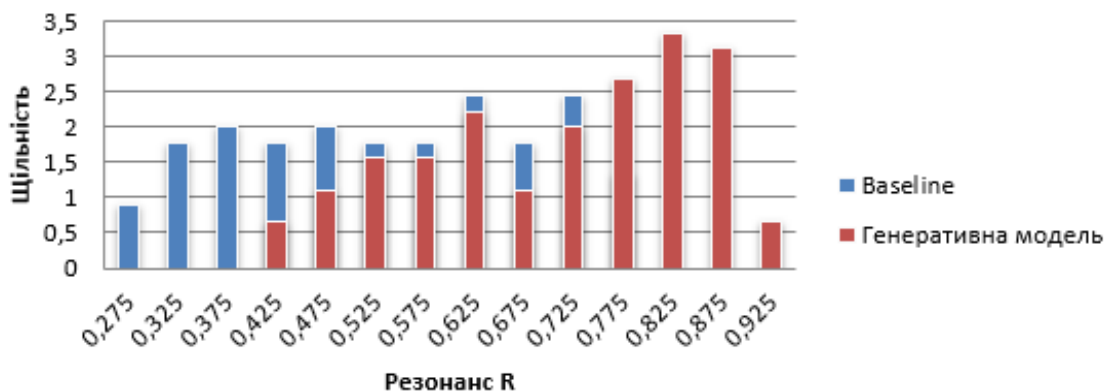


Рис. 1. Порівняння розподілу резонансу

На графіку порівняння розподілу показників резонансу для базової (baseline) моделі та генеративної моделі помітно, що базова модель (baseline) має значну концентрацію значень у низькому та середньому діапазонах (0.3–0.6), причому частина сценаріїв потрапляє навіть у зону нижче 0.4. Натомість генеративна модель демонструє зсув розподілу вправо: більшість її значень зосереджена в області високого резонансу (0.75–0.88), а низькорезонансні сценарії практично відсутні. Така різниця у формах розподілів свідчить про підвищення стабільності та якості сценаріїв, сформованих генеративною моделлю. Крім того, щільність розподілу у верхніх діапазонах значно більша саме для генеративної моделі, що підтверджує її здатність частіше генерувати високорезонансні відповіді.

Таблиця 1

Порівняння baseline-моделі та моделі з генеративною моделлю

Показник	Базова (baseline) модель	Генеративна модель	Зміна
Середній резонанс R	0.62	0.71	+14.5%
Частка ефективних сценаріїв ($R > 0.75$)	28%	43%	+15%
Частка сценаріїв з низьким резонансом ($R < 0.4$)	22%	11%	–11%
Конверсія	12%	13.4%	+11.7%

Результати демонструють, що генеративний підхід суттєво підвищує узгодженість створюваних відповідей із поточним станом клієнта. Середній резонанс R зріс із 0.62 до 0.71, що становить приблизно 14.5% покращення. Частка сценаріїв із високим резонансом ($R > 0.75$) зросла з 28% до 43%, що підтверджує здатність моделі генерувати більш адаптивні та релевантні комунікаційні варіанти. Водночас частка сценаріїв із низьким резонансом ($R < 0.4$) зменшилася удвічі – із 22% до 11%, що свідчить про підвищення стабільності генеративного підходу та зменшення ризику появи нерелевантних або некоректних відповідей. Оцінювана конверсія демонструє потенційний приріст на рівні близько 11.7%, що узгоджується з підвищенням резонансності сценаріїв. У цілому проведене дослідження підтверджує, що генеративна модель здатна значно підвищити якість персоналізованої взаємодії й забезпечити більш точне пристосування відповідей до емоційного та когнітивного контексту клієнта. Це робить її перспективним інструментом для подальших досліджень та інтеграції у реальні CRM-системи.

Висновки

У роботі було запропоновано підхід до формування резонансних сценаріїв комунікації у CRM-системах на основі генеративних мовних моделей із урахуванням емоційного та когнітивного стану клієнта. На

відміну від традиційних заснованих на правилах рішень та моделей, орієнтованих виключно на аналіз тональності тексту, представлена концепція ґрунтується на двовекторному описі користувача ($E_s + C_s$), що забезпечує більш глибоке та точне розуміння його намірів, емоційної динаміки, рівня залученості та інформаційної потреби. Такий підхід дозволив сформувати механізм генерації персоналізованих сценаріїв, здатних адаптивно змінювати зміст, тональність і комунікаційну стратегію відповідно до поточного профілю клієнта.

Експериментальні результати підтвердили ефективність запропонованої моделі порівняно з шаблонним базовим сценарієм. Зокрема, зафіксовано зростання середнього показника резонансності відповіді, збільшення частки високоякісних сценаріїв із високим значенням R_s та зменшення кількості нерелевантних варіантів. Крім того, спостерігається покращення конверсійних метрик взаємодії, що свідчить про практичну доцільність впровадження підходу в комерційні CRM-платформи.

Наукова новизна роботи полягає у формалізації резонансно-орієнтованої моделі генеративної комунікації в CRM, яка поєднує двовекторне представлення емоційного та когнітивного стану клієнта з адаптивною генерацією сценаріїв за допомогою LLM. Запропоновано резонансну функцію R_s як метрику відповідності між станом користувача та характеристиками відповіді, а також динамічну структуру промпта, що враховує психологічний профіль, історію діалогу та бізнес-правила. Це створює основу для розроблення інтелектуальних CRM-систем, здатних до гнучкої, контекстно-адаптивної і психологічно орієнтованої комунікації з користувачем.

Подальші дослідження доцільно спрямувати на розширення набору когнітивних та емоційних показників, впровадження багатомодальних сигналів (голос, поведінкові патерни, реакції у реальному часі), а також проведення експериментів із реальними користувачами та живими бізнес-сценаріями. Перспективним напрямом є також оптимізація RL-компонентів для підвищення стабільності адаптації моделі та розробка механізмів етичного контролю генеративних сценаріїв.

Література

1. Zhou M., Duan N., Liu S., Shum H. Y. Progress in neural NLP: modeling, learning, and reasoning // *Engineering*. – 2020. – Vol. 6, № 3. – С. 275–290.
2. Kolomeitsev K. LLM Modules: Knowledge Transfer from a Large to a Small Model using Enhanced Cross-Attention // 2025. – С. 3.
3. Sharifani K., Amini M. Machine learning and deep learning: A review of methods and applications // *World Information Technology and Engineering Journal*. – 2023. – Vol. 10, № 07. – С. 3897.
4. Dam S. K., Hong C. S., Qiao Y., Zhang C. A complete survey on llm-based ai chatbots // 2024. – С. 5.
5. Kim J. K., Chua M., Rickard M., Lorenzo A. ChatGPT and large language model (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine // *Journal of Pediatric Urology*. – 2023. – Vol. 19, № 5. – С. 598.
6. Гордеев А. Основні поняття машинного навчання з підкріпленням. – 2024. – С. 358.
7. Claassen J., Höhne J. K., Bach R., Haensch A. C. Identifying bots through LLM-generated text in open narrative responses: A proof-of-concept study. – С. 2
8. Marvin G., Hellen N., Jjingo D., Nakatumba-Nabende J. Prompt engineering in large language models // *International conference on data intelligence and cognitive informatics*. – 2023. – С. 387.
9. Sarker I. H., Janicke H., Ferrag M. A., Abuadba A. Multi-aspect rule-based AI: Methods, taxonomy, challenges and directions towards automation, intelligence and transparent cybersecurity modeling for critical infrastructures // *Internet of Things*. – 2024. – С. 25.
10. Mu T., Helyar A., Heidecke J., Achiam J., Vallone A., Kivlichan I., Weng L. Rule based rewards for language model safety // *Advances in Neural Information Processing Systems*. – 2024. – С. 37.

References

1. Zhou, M., Duan, N., Liu, S., & Shum, H. Y. (2020). Progress in neural NLP: modeling, learning, and reasoning. *Engineering*, 6(3), 275–290.
2. Kolomeitsev, K. (2025). LLM Modules: Knowledge Transfer from a Large to a Small Model using Enhanced Cross-Attention, 3.
3. Sharifani, K., & Amini, M. (2023). Machine learning and deep learning: A review of methods and applications. *World Information Technology and Engineering Journal*, 10(07), 3897.
4. Dam, S. K., Hong, C. S., Qiao, Y., & Zhang, C. (2024). A complete survey on LLM-based AI chatbots, 5.
5. Kim, J. K., Chua, M., Rickard, M., & Lorenzo, A. (2023). ChatGPT and large language model (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine. *Journal of Pediatric Urology*, 19(5), 598.
6. Gordieiev, A. (2024). Osnovni poniattia mashynnoho navchannia z pidkriplenniam [Basic concepts of reinforcement learning]. 358.
7. Claassen, J., Höhne, J. K., Bach, R., & Haensch, A. C. (2024). Identifying bots through LLM-generated text in open narrative responses: A proof-of-concept study, 2.
8. Marvin, G., Hellen, N., Jjingo, D., & Nakatumba-Nabende, J. (2023). Prompt engineering in large language models. *International Conference on Data Intelligence and Cognitive Informatics*, 387.
9. Sarker, I. H., Janicke, H., Ferrag, M. A., & Abuadba, A. (2024). Multi-aspect rule-based AI: Methods, taxonomy, challenges and directions towards automation, intelligence and transparent cybersecurity modeling for critical infrastructures. *Internet of Things*, 25.
10. Mu, T., Helyar, A., Heidecke, J., Achiam, J., Vallone, A., Kivlichan, I., & Weng, L. (2024). Rule based rewards for language model safety. *Advances in Neural Information Processing Systems*, 37.