УДК 004

РОЗРОБКА ІНФОРМАЦІЙНОЇ СИСТЕМИ СТАТИСТИЧНОГО АНАЛІЗУ ДЛЯ ЧАТ-БОТ ПЛАТФОРМ

Кам'янский Д.В.[1]
[1]Одеський національний морський університет
[1]Україна, Одеса
n.o.novikova25@gmail.com

Анотація. *В матеріалі розглянуто проблематику відсутності системної аналітики в процесі функціонування чат-бот платформ, що обмежує можливості оптимізації сценаріїв спілкування та підвищення ефективності користувацької взаємодії. Запропоновано концепцію розробки інформаційної системи статистичного аналізу, орієнтованої на автоматизоване збирання, обробку та інтерпретацію діалогових даних. Описано архітектуру, програмні компоненти та потенціал практичного застосування рішення у сфері цифрових комунікацій і бізнес-аналітики.*
Ключові слова: *чат-бот; інформаційна система; статистичний аналіз; аналітика даних; ефективність взаємодії.*

Вступна частина. Сучасні інформаційні технології дедалі активніше інтегруються у комунікаційні процеси між бізнесом, державними структурами та користувачами. Значну роль у цьому відіграють чат-боти — автоматизовані програмні агенти, які забезпечують обмін повідомленнями у режимі реального часу та виконують низку функцій від інформування до обробки запитів і підтримки клієнтів. Популярність чат-ботів у таких платформах, як Telegram, Facebook Messenger, WhatsApp чи корпоративні інтранет-системи, зумовила швидке зростання обсягів діалогових даних, що утворюються під час користувацької взаємодії [1,2].

Разом із цим зростає потреба у побудові ефективних механізмів статистичного аналізу цих даних. Більшість розробників чат-ботів обмежуються базовою аналітикою — кількістю запитів, кількістю унікальних користувачів, середнім часом відповіді. Такі показники не відображають повною мірою реальної ефективності сценаріїв спілкування або ступеня задоволеності користувача. Крім того, у багатьох випадках відсутній інструментарій для порівняння різних версій діалогових моделей чи відстеження динаміки поведінкових змін аудиторії [3,4].

Зважаючи на це, постає науково-практична задача створення інформаційної системи, здатної здійснювати глибокий статистичний аналіз чат-бот трафіку, виявляти закономірності, прогнозувати ефективність і формувати рекомендації для вдосконалення роботи ботів. Така система повинна бути інтегрованою, масштабованою й орієнтованою на роботу з великими обсягами неструктурованих текстових даних.

Метою дослідження є розробка концепції, архітектури та методів реалізації інформаційної системи статистичного аналізу для чат-бот платформ, яка забезпечує збирання, обробку, збереження та інтерпретацію даних користувацьких діалогів. Додатково метою є обґрунтування технічних і програмних засобів для побудови модульного рішення, що здатне підтримувати різні платформи та забезпечувати глибоку аналітику без втручання користувача в процеси обробки.

Основна частина. З огляду на значну кількість чат-ботів у комерційних, освітніх і державних сервісах, виникає необхідність формування аналітичних механізмів, які дозволяють досліджувати статистичні закономірності діалогів. Ключова проблема полягає у розрізненості джерел даних, високій динамічності повідомлень, різних форматах логів та складності подальшої агрегації інформації. Існуючі аналітичні платформи здебільшого не адаптовані до контексту чат-ботів, оскільки не враховують структурні елементи діалогу, послідовність повідомлень, тематику та результативність взаємодії.

Традиційні інструменти BI-аналітики не дозволяють автоматично проводити глибинний статистичний аналіз на рівні діалогових сутностей. У результаті власники ботів часто приймають рішення інтуїтивно, без опори на достовірні дані. Це призводить до неефективного використання обчислювальних ресурсів, зниження продуктивності ботів і втрачених можливостей у взаємодії з користувачами.

Розроблювана система призначена для вирішення зазначених проблем шляхом створення єдиного аналітичного середовища, яке забезпечує наскрізний процес обробки даних — від збору діалогів до отримання статистичних показників та рекомендацій. Архітектура будується на мікросервісній моделі, що забезпечує масштабованість і незалежність компонентів.

Основу рішення складає тришаровий підхід: рівень даних, аналітичний рівень і рівень представлення. На рівні даних відбувається інтеграція з чат-бот платформами через API або webhook-з'єднання. Дані про повідомлення користувачів, час запитів, ключові слова, реакції та контексти діалогів надходять у проміжне сховище. Для уніфікації застосовується ETL-процес, який здійснює вилучення, трансформацію та завантаження даних у центральну базу.

Для обробки даних обрано PostgreSQL із розширенням TimescaleDB, що оптимізує роботу з часовими рядами. На цьому етапі реалізується очистка, токенізація тексту, виділення частоти повідомлень, середнього часу відповіді та коефіцієнтів переходу між станами діалогу.

Аналітичний рівень реалізує модулі обчислення статистичних метрик, кластеризації діалогів та побудови кореляцій між параметрами поведінки користувачів. Для цього застосовуються алгоритми машинного навчання — зокрема, методи k-means та дерева рішень для визначення типових моделей спілкування.

Рівень представлення відповідає за візуалізацію результатів. Користувацький інтерфейс створюється на базі React, а графічна інтерпретація аналітичних показників реалізується за допомогою бібліотек Chart.js і D3.js. Завдяки цьому користувач отримує інтерактивні панелі з можливістю фільтрації, порівняння та експорту результатів у різних форматах.

На серверному рівні використовується Python із фреймворком FastAPI для створення REST API, що обслуговує запити до системи. Асинхронна обробка даних забезпечується за допомогою брокера повідомлень RabbitMQ, що дозволяє працювати із потоковими подіями. Для масштабування застосовується контейнеризація через Docker.

Наукова новизна розробки полягає у формуванні методичного підходу до статистичного аналізу діалогових даних, який об'єднує принципи класичної статистики з алгоритмами машинного навчання для глибшого розуміння поведінки користувачів. Система дозволяє переходити від поверхневих метрик до багатовимірних моделей оцінки ефективності комунікації, що є новим напрямом у дослідженнях взаємодії людини та штучного інтелекту.

Практична значущість рішення полягає у можливості його інтеграції у будь-які існуючі чат-бот платформи. Це відкриває перспективи використання у бізнесі для моніторингу клієнтських запитів, у сфері освіти для аналізу ефективності навчальних ботів, а також у державному секторі для оцінки зворотного зв'язку з громадянами.

Висновки. Проведене дослідження підтвердило доцільність створення інформаційної системи статистичного аналізу для чат-бот платформ. Запропонована архітектура поєднує сучасні технологічні засоби, гнучку модульну структуру та аналітичні механізми, здатні перетворювати великі обсяги діалогових даних у практично корисну інформацію. Така система сприятиме підвищенню ефективності взаємодії користувачів із ботами, оптимізації діалогових сценаріїв та обґрунтованому прийняттю управлінських рішень.

Подальші дослідження доцільно спрямувати на вдосконалення модулів інтелектуального аналізу з використанням моделей обробки природної мови (NLP) та глибинного навчання, що дозволить реалізувати семантичну інтерпретацію діалогів і рекомендаційні системи нового покоління.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Aggarwal, C.C. (2015). *Data mining: The textbook*. Springer.
[2] Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach*. Pearson.

[3] Chollet, F. (2021). *Deep learning with Python* (2nd ed.). Manning Publications.
[4] Sharda, R., Delen, D., & Turban, E. (2023). *Business intelligence and analytics: Systems for decision support*. Pearson.

Kamyansky D.V.

Development of an information system for statistical analysis for chat-bot platforms

Annotation. *The paper examines the problem of the lack of systematic analytics in chatbot platforms, which limits the optimization of conversational scenarios and user engagement efficiency. A concept for developing an information system for statistical analysis focused on automated data collection, processing, and interpretation is proposed. The architecture, software components, and practical applications in digital communication and business analytics are discussed.*

Keywords: *chatbot; information system; statistical analysis; data analytics; interaction efficiency.*