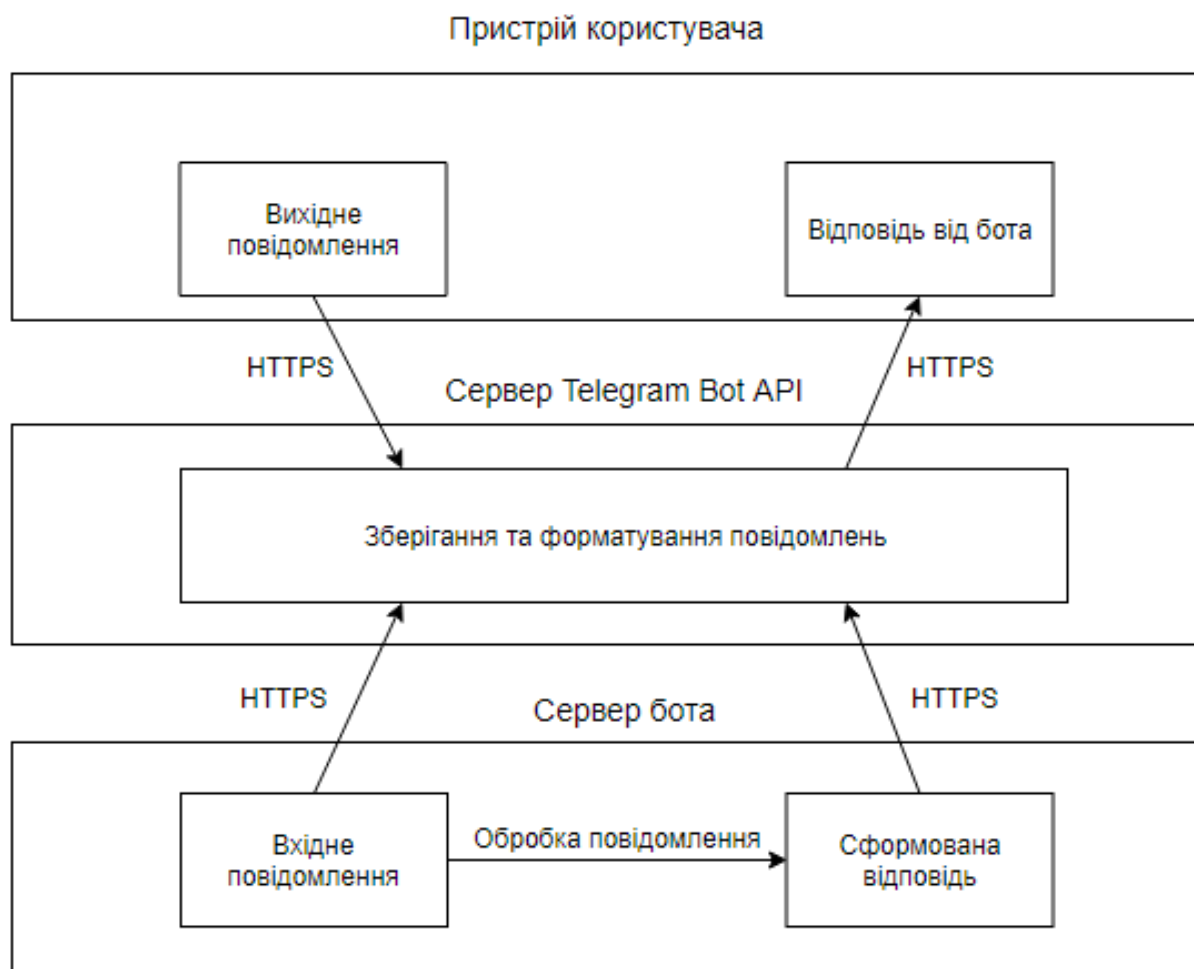
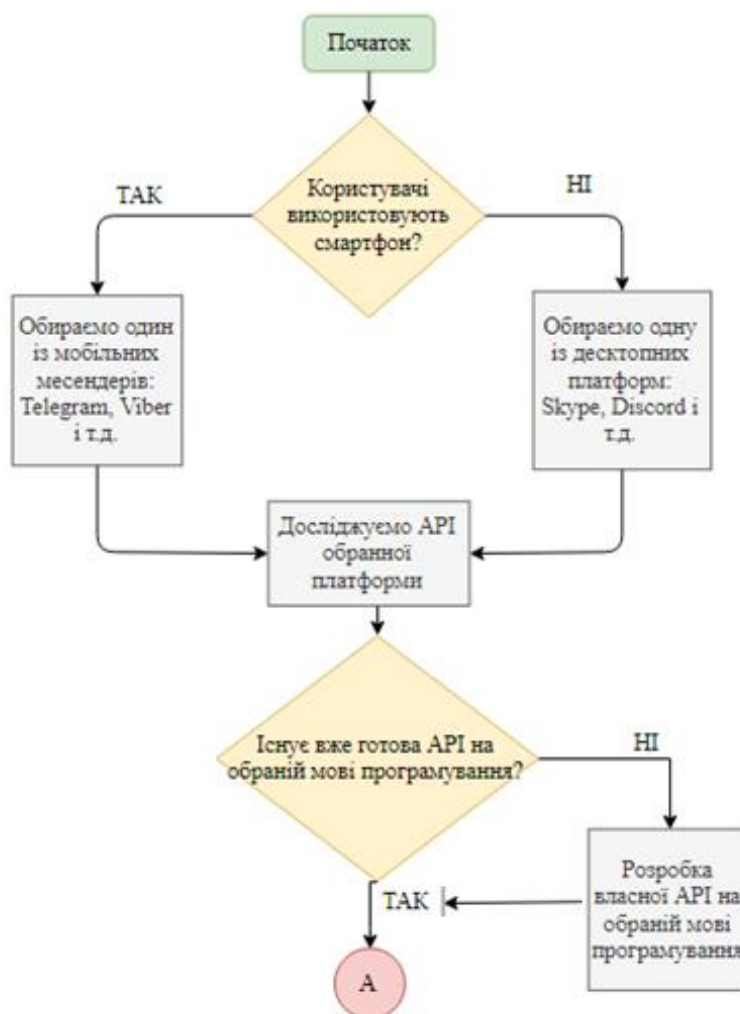


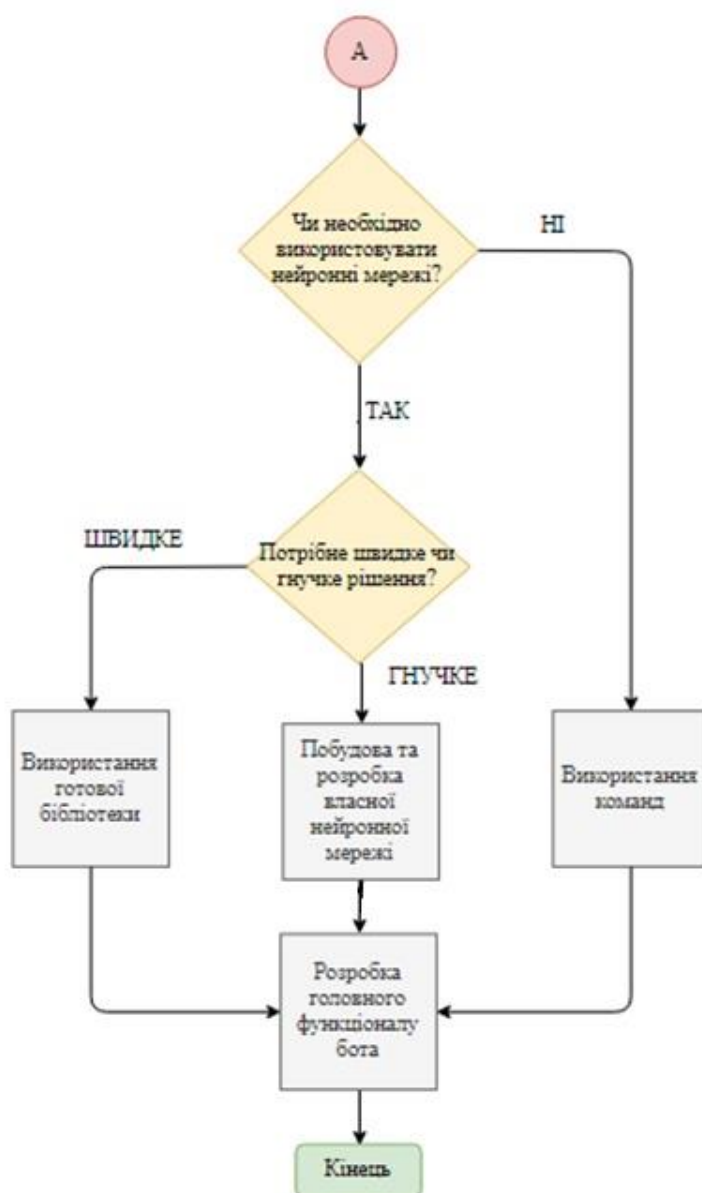
ДОДАТОК А
Схема елементів Telegram з використанням бота



ДОДАТОК Б

Алгоритм розробки бота





ДОДАТОК В
Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук

Кафедра Інформаційних управляючих систем

МАГІСТЕРСЬКА АТЕСТАЦІЙНА РОБОТА
ГРАФІЧНИЙ МАТЕРІАЛ

Дослідження методів розробки бот-систем
(тема роботи)

Магістрант гр. ІУСТм-18-1
(шифр групи)

(підпис)

Воловик С.М.
(прізвище, ініціали)

Науковий керівник роботи

(підпис)

проф. Левикін В.М.
(посада, ініціали, прізвище)

2019 р.

1 ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Таблиця 1.1 – Загальна характеристика роботи

Тема МАР	Дослідження методів розробки бот-систем
Актуальність	Велику кількість часу люди проводять, набираючи текст в різних месенджерах (будь то Telegram, Facebook, Slack і т.д) робить ці самі месенджери швидко зростаючим сегментом на ринку додатків. Люди використовують їх не тільки для дружніх листувань, але і в інтересах розвитку бізнесу. Швидкий розвиток технологій штучного інтелекту, машинного навчання дозволяє сучасним чат-ботам поводитися максимально наближено до того, як поведуться люди.
Мета досліджень	Дослідження популярності сучасних платформ, месенджерів, визначення програмної платформи (фреймворк) та інтерфейс прикладних рішень і бібліотек API з великою кількістю доступних інструментів (функцій, класів, структур), які підходять для сучасної розробки бота.
Задачі досліджень	Дослідити існуючі підходи для розробки бота; Дослідити існуючі проблеми чат-ботів; Дослідити методи машинного навчання для вдосконалення ботів, та їх взаємодію з користувачем; Розробка прикладної моделі.
Об'єкт та предмет дослідження	Об'єктом дослідження є процес розробки бот-систем. Предмет дослідження – методи розробки бот-систем.
Наукова новизна роботи	Новизна дослідження полягає в дослідженні методів та рекомендації використання методів машинного навчання при розробці чат-бота з метою персоналізації та підвищення якості обслуговування користувачів.

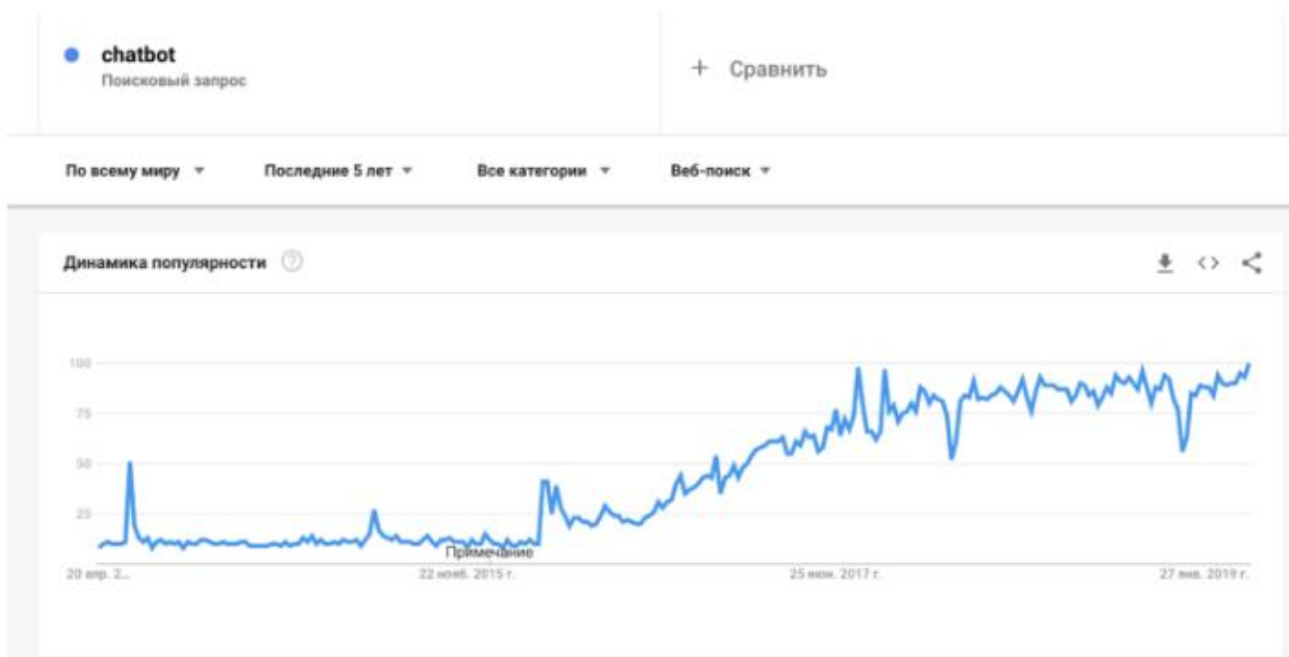


Рисунок 1.1 – Динаміка популярності чат-ботів на сьогодні (Джерело: Google Trends)

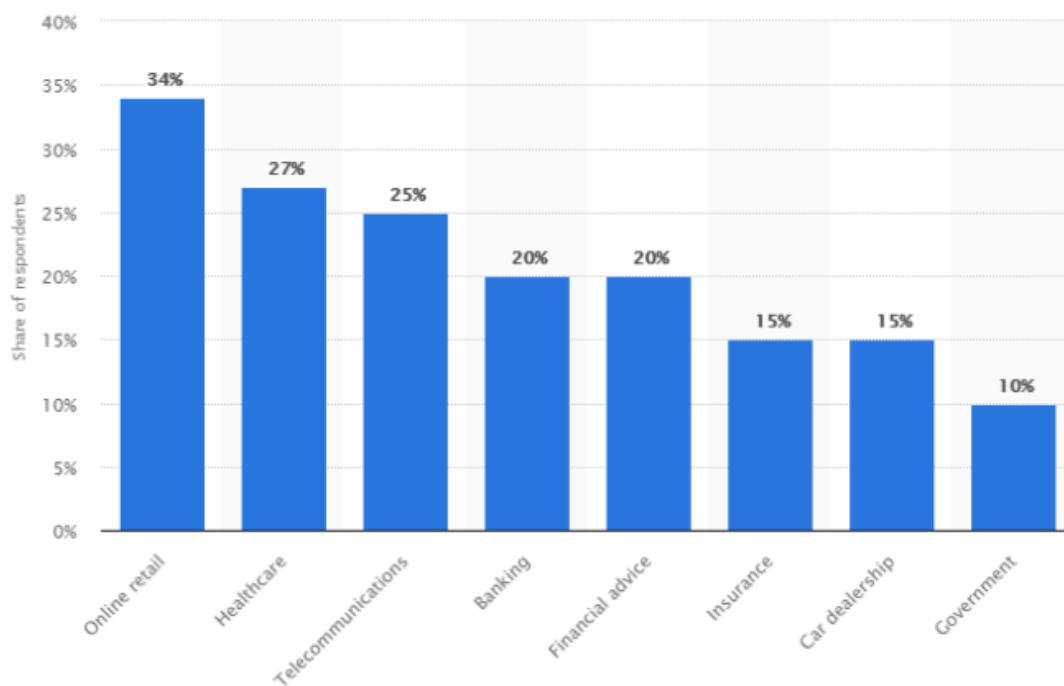


Рисунок 1.2 – Данні опитування Statista

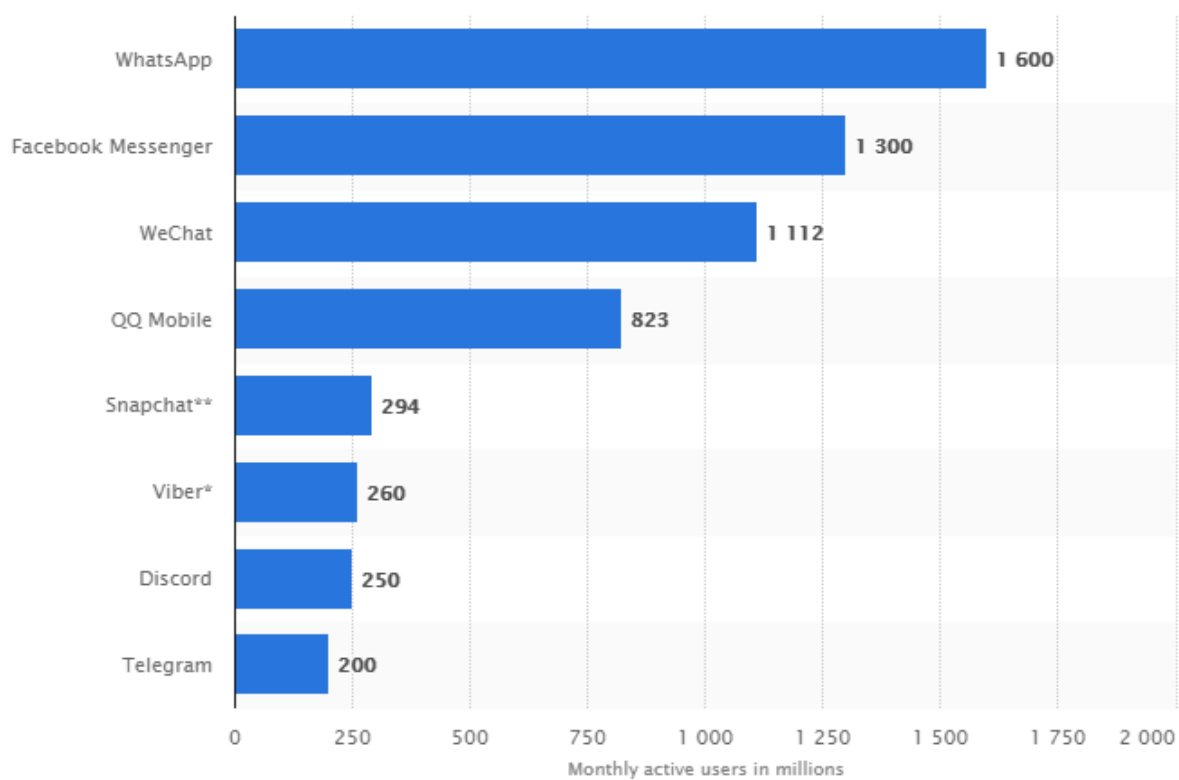


Рисунок 1.3 – Кількість користувачів в месенджерах за місяць

2 КЛАСИФІКАЦІЙ ТА МОЖЛИВІ ВАРІАНТІВ РОЗРОБКИ ЧАТ-БОТІВ

Таблиця 2.1 – Класифікація чат-ботів

Бот-автоматизатор	Виконує за користувача кілька рутинних дій в автоматичному режимі, економить йому час. Корисний для таких завдань, як відстеження посилок або розкладу транспорту, замовлення таксі, автоматична відправка листів за результатами подій в CRM
Бот-помічник	Вміє розпізнавати запити від клієнтів і давати на них відповіді, чи не залучаючи людини в процес. Якщо впирається в стелю своїх знань, перемикає клієнта на менеджера. Менеджер в свою чергу отримує повну історію звернення – з озвученими клієнтом персональними даними, питанням, історією спілкування. Такий бот «розвантажує» менеджерів, позбавляючи від рутини. Самонавчального, тому з часом може закривати все більше складних завдань. Може використовуватися, наприклад, службами підтримки для відповіді на найпопулярніші запити.
Вбудований бот	Бот з елементами ШІ, йому можна надати великий обсяг даних, в якому він буде шукати інформацію краще і швидше людини. Корисний там, де потрібно орієнтуватися у величезних обсягах даних, наприклад, в документах страхових або юридичних фірм
Бот, який допомагає оформити замовлення	Цей бот може відшукати потрібний товар і оформити замовлення через інтерфейс месенджера швидше, ніж це можна було б зробити через сайт. Економить час користувача за рахунок того, що виконує максимальну кількість завдань через інтерфейс чату, і перекидає в браузер для завершення оформлення замовлення (авторизації або оплати) автоматично, позбавляючи від необхідності щось шукати вручну. Корисний для сервісів бронювання житла, квитків і т.д.
Бот-збирач лідів	Збирає дані про клієнтів, зацікавлених в замовленні, уточнює у них персональні дані та інформацію про цікавить товар / послугу. Також береже менеджерів компанії від рутинної роботи. Корисний всюди, де потрібно збирати досить однотипні дані в користувачів.

Таблиця 2.2 – Сервіси та фреймворки для створення чат-ботів та їх загальні недоліки

Назва	Недоліки
<p>Сервіси</p> <ul style="list-style-type: none"> • Flow XO • Chatfuel • Dialogflow 	<p>Неможливо заблокувати відповідність наміру, якщо присутній контекст.</p> <p>Навчальний розділ все ще знаходиться в бета-версії.</p> <p>Навіть якщо історії є потужною концепцією, є випадки, коли важко контролювати потік розмови, а бот схильний неправильно розуміти запити користувачів.</p>
<p>Фреймворки:</p> <ul style="list-style-type: none"> • Errbot (Python) • Hubot (CoffeeScript) 	<p>Є складні системи фільтрів та налаштувань.</p>

А для створення ботів без використання сервісу, використовують API платформи. Наприклад, Bot API представляє собою HTTP-інтерфейс для роботи з ботами в Telegram. Кожен бот – це спеціальний аккаунт, створений для автоматичного оброблення та відправлення повідомлень.

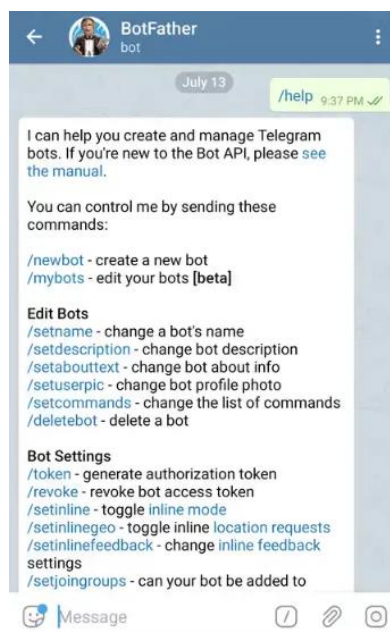


Рисунок 2.1 – Результат обробки команди «/help» ботом BotFather

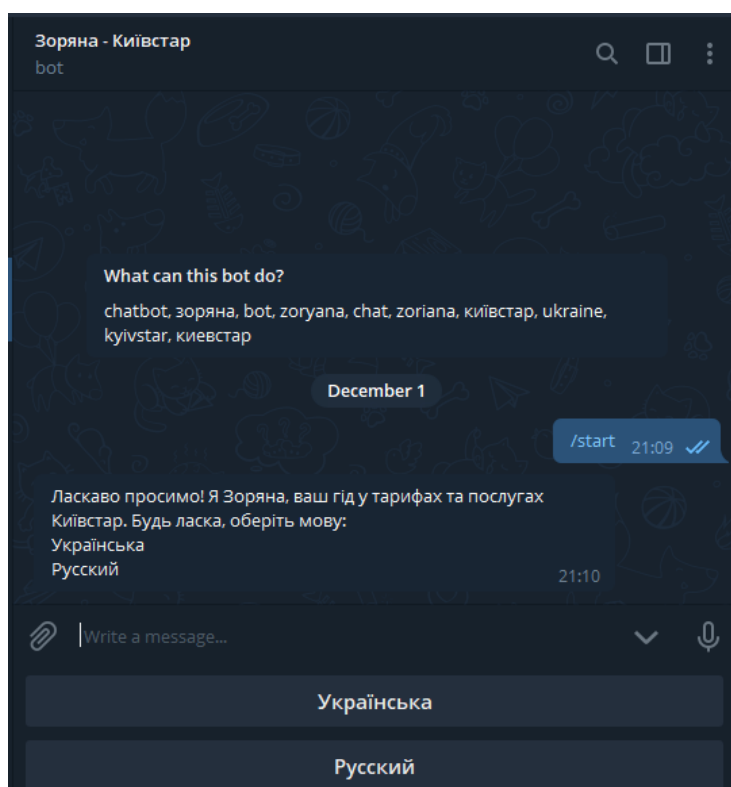


Рисунок 2.2 – Інтерфейс бота Зоряни компанії «Київстар» в Telegram

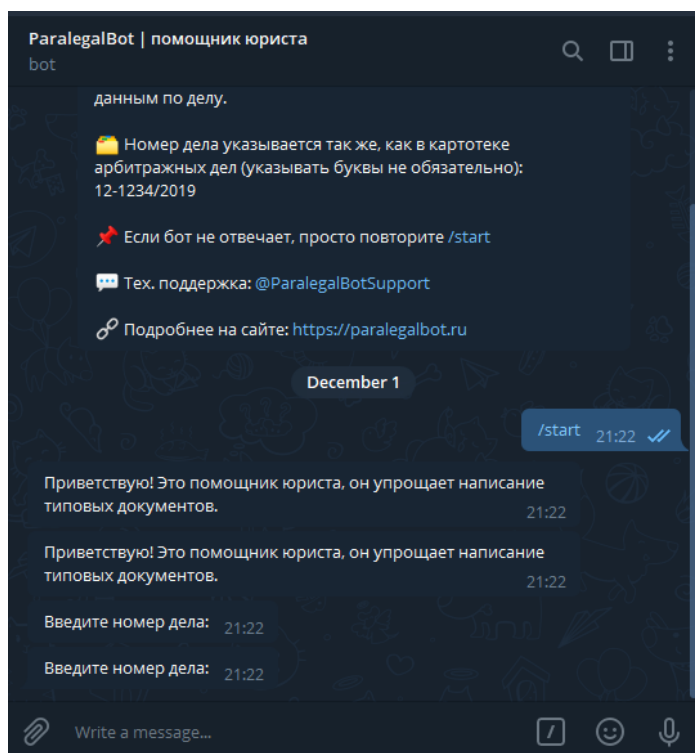


Рисунок 2.3 – Интерфейс бота Paralegalbot в Telegram

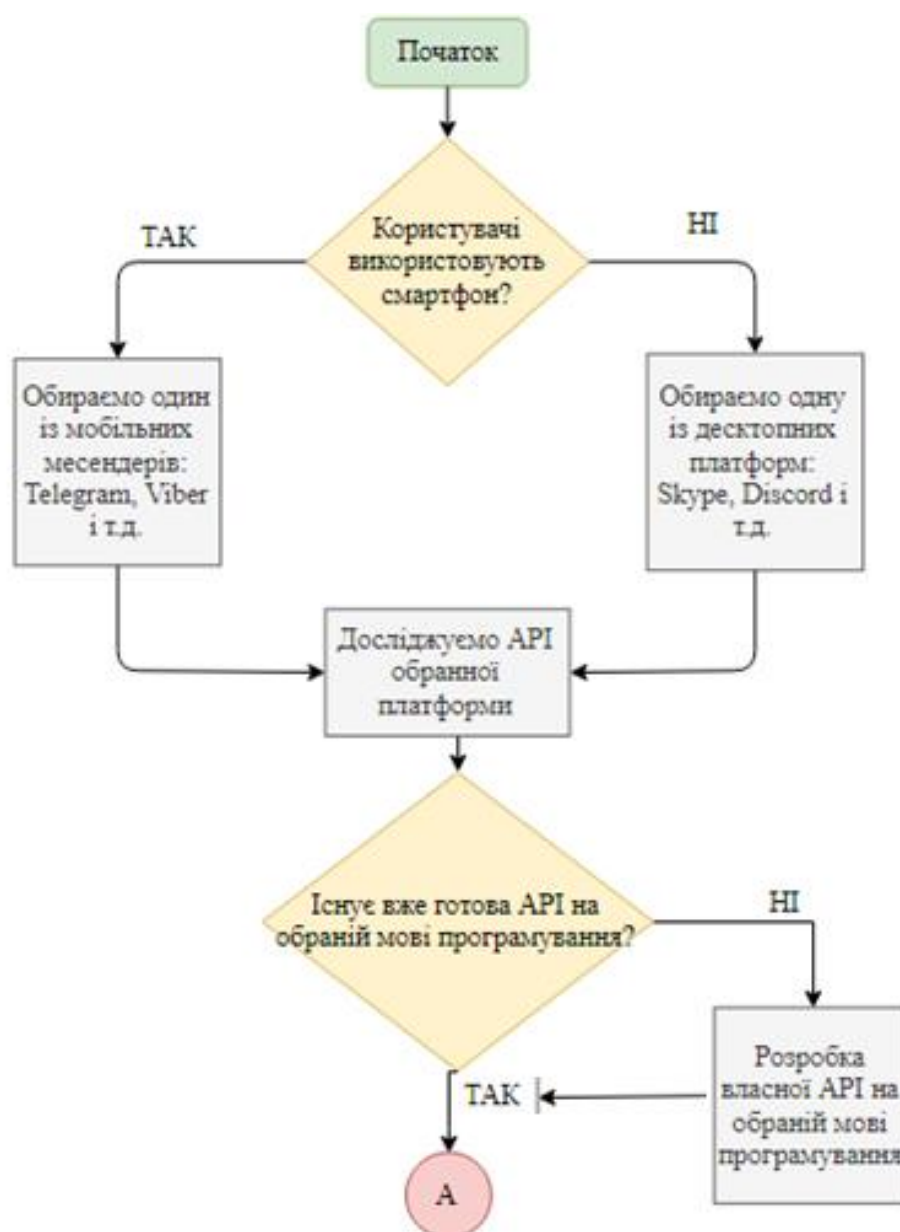


Рисунок 2.4 – Алгоритм розробки бота (частина 1)

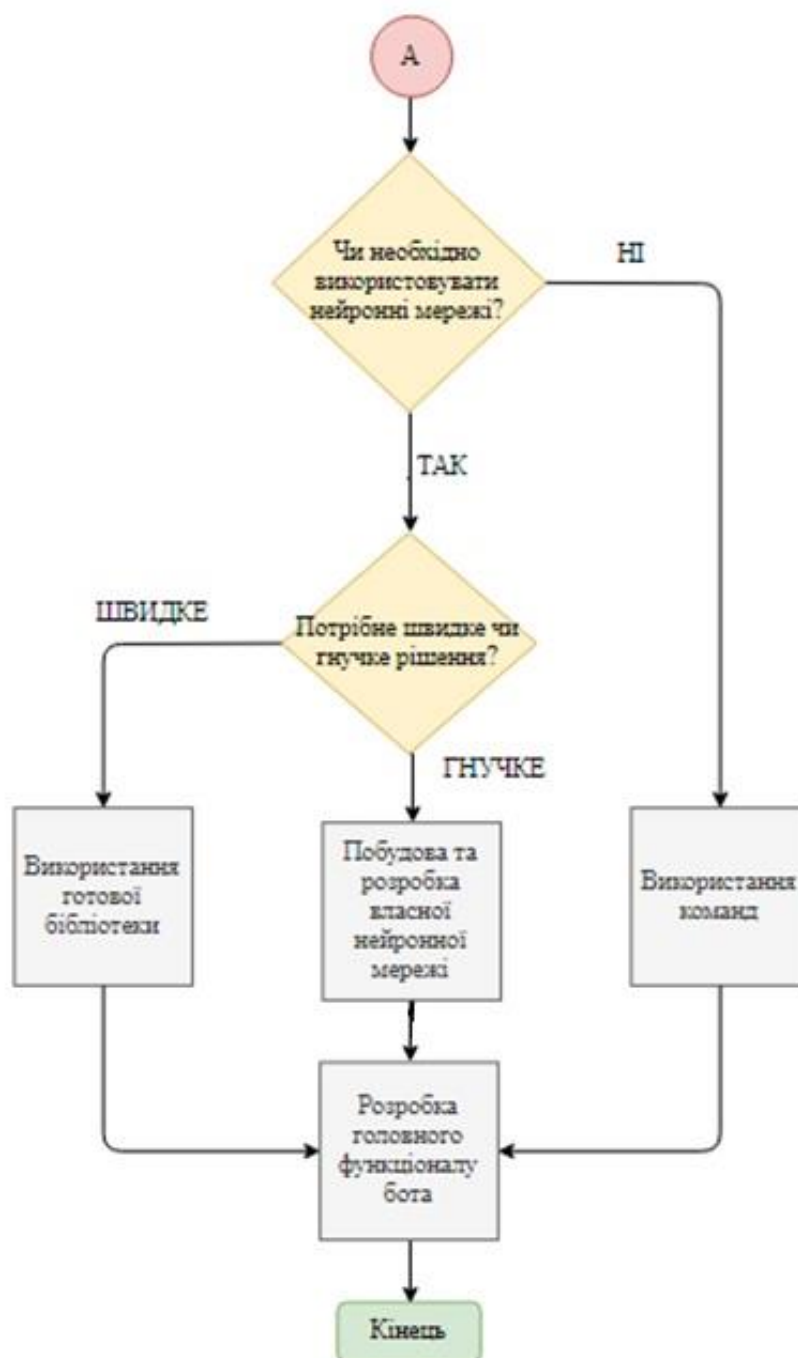


Рисунок 2.5 – Алгоритм розробки бота (частина 2)

3 МЕТОДИ МАШИННОГО НАВЧАННЯ, ЇХ ЗАСТОСУВАННЯ В БОТ РОЗРОБЦІ

Таблиця 3.1 – Загальна характеристика методів машинного навчання

<p>Послідовність до послідовності (Seq2Seq)</p>	<p>Це найбільш часто використовувана архітектура в машинному перекладі та системах зіставлення питань нейронної мережі.</p> <p>1) Енкодер (enc) послідовно зчитує вхідний пропозицію і видає вектор контексту «C». На малюнку передачі вектора контексту від енкодера до декодера відповідає стрілка $h_4 \rightarrow dec_1$.</p> <p>2) Отримавши опис вхідної пропозиції в вигляді «C», декодер (dec) генерує слова на іншій мові послідовно. Окремим «словом» в словнику мови Б позначимо початок пропозиції «BOS» і кінець пропозиції «EOS».</p> <p>Декодер можна представити у вигляді функції:</p> $h_t, p_t = dec(h_{t-1}, w_{t-1})$ <p>Тут h_t - вектор прихованого стану декодера на кроці t, p_t - розподіл за словами на етапі t, $w_{(t-1)}$ це слово, яке вибрано в якості основної гіпотези виходячи з розподілу $p_{(t-1)}$ на етапі «$t-1$».</p> <p>Для енкодера аналогічне рівняння виглядає наступним чином:</p> $h_t = enc(h_{t-1}, w_t)$
---	---

<p>Слово у вектор (Word2Vec)</p>	<p>Ідея цілком очевидна з назви: уявити слово у вигляді вектора з координатами (x_1, x_2, \dots, x_n).</p> <p>Мета та корисність Word2vec – згрупувати вектори подібних слів разом у векторі простору. Тобто виявляє схожість математично. Word2vec створює вектори, які розподіляють числові уявлення про ознаки слова, такі функції, як контекст окремих слів. Це робиться без втручання людини.</p> <p>Вихід нейронної мережі Word2vec – це словниковий запас, в якому кожен елемент має приєднаний до нього вектор, який можна подати в мережу глибокого вивчення або просто запитати для виявлення зв'язків між словами.</p> <p>Недоліком word2vec є те, що з його допомогою слова, які не зустрічаються у навчальному наборі, не можуть бути представлені.</p>
<p>Термін Частота /Зворотна частота (TF-IDF)</p>	<p>Цей процес включає обчислення відносної ваги кожного слова протягом усієї вибірки приклад на рисунку 3.5. Вага слова обчислюється діленням частоти зустрічі слова на частоту його зустрічі в інших питаннях. Таким чином, слова, що означають послуги, технології та галузі, що цікавлять користувачів, мають масштаби, які набагато більше, ніж слова, що складають решту питань. Потім на основі підготовленого векторизатора обробляється вся база запитань та навчання бота.</p>

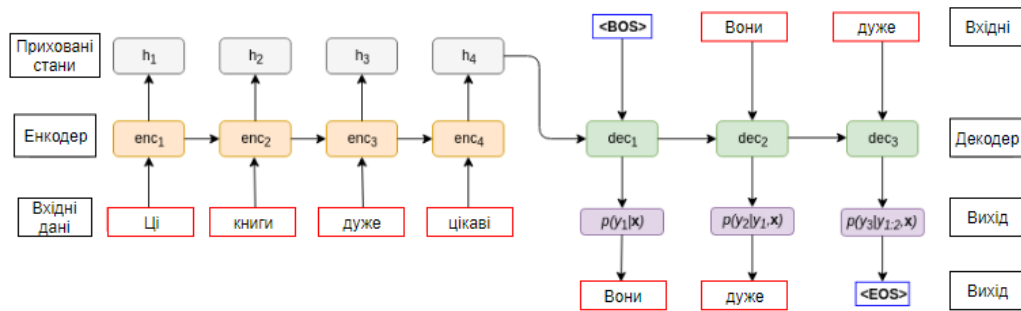


Рисунок 3.1 – Схема моделі перекладу

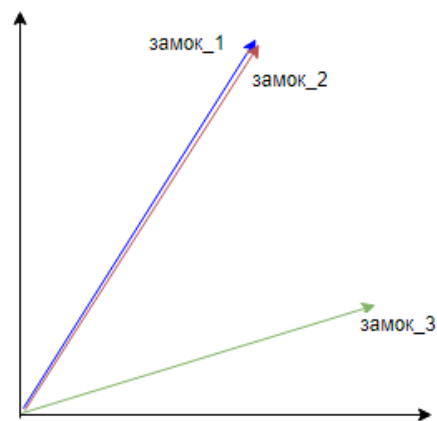


Рисунок 3.2 – Вектори схожих слів «замок_1», «замок_2» та омоніма «замок_3»

купити історичну книгу

word 1 word 2 word 3

	word 1	word 2	word 3
query	1	1	1
doc 1	1	1	
doc 2	2	2	1
doc 3		5	5
doc 4		1	

Рисунок 3.3 – Приклад відносної ваги слів (вектор TF-IDF)

Таблиця 3.2– Точність методу TF-IDF

	~ 10 ключових слів	~ 30 ключових слів
ML + tf-idf	0,92	0,85

4 ПРИКЛАДНА ПРОГРАМА COFFEE BAR ЧАТ-БОТА НА ОСНОВІ РОЗПІЗНАВАННЯ КЛЮЧОВИХ СЛІВ, МЕНЮ / КНОПОК ТА КОНТЕКСТУ

```
var http = require("https");

var options = {
  "method": "GET",
  "hostname": "api.chatbot.com",
  "port": null,
  "path": "/stories",
  "headers": {
    "authorization": "Bearer ${DEVELOPER_ACCESS_TOKEN}"
  }
};

var req = http.request(options, function (res) {
  var chunks = [];

  res.on("data", function (chunk) {
    chunks.push(chunk);
  });

  res.on("end", function () {
    var body = Buffer.concat(chunks);
    console.log(body.toString());
  });
});

req.end();
```

Рисунок 4.1 – Приклад запиту аутентифікації (код Node js)

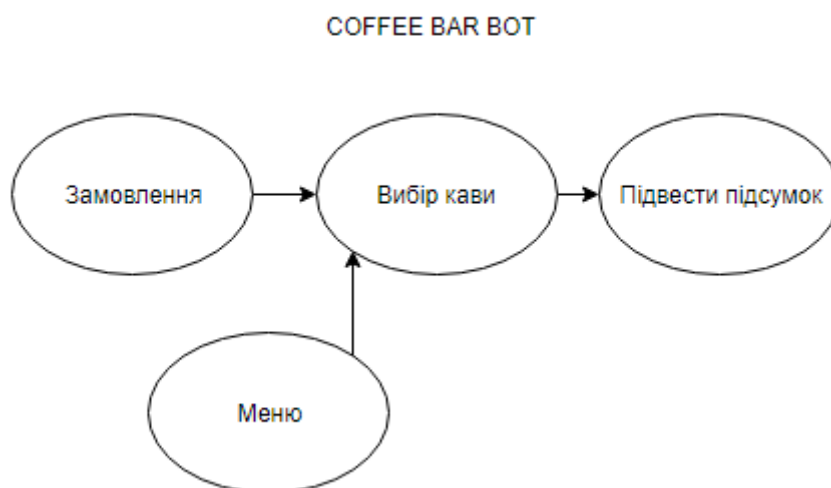


Рисунок 4.2 – План побудови чат-бота COFFEE BAR

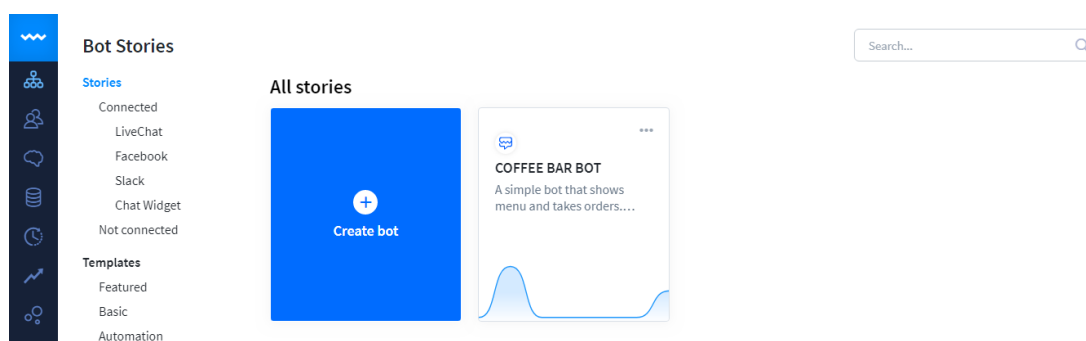


Рисунок 4.3 – Інформаційна панель API ChatBot

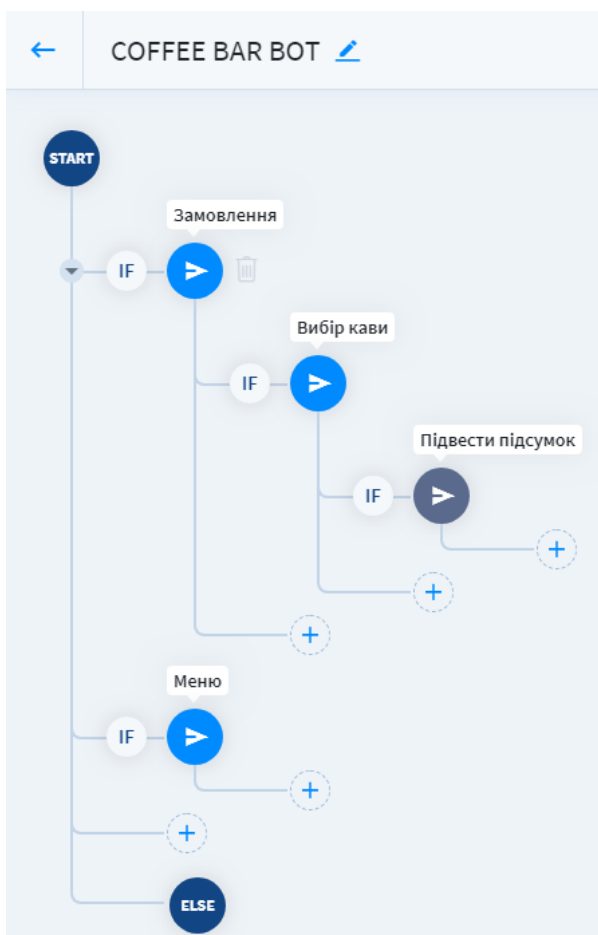


Рисунок 4.4 – Історія умов програми COFFEE BAR BOT

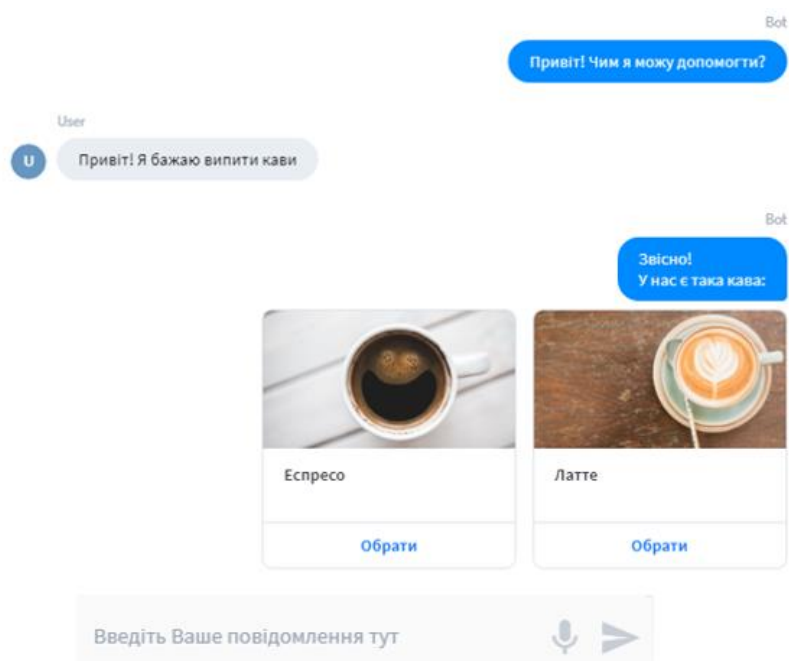


Рисунок 4.5 – Прикладна програма COFFEE BAR BOT в API ChatBot

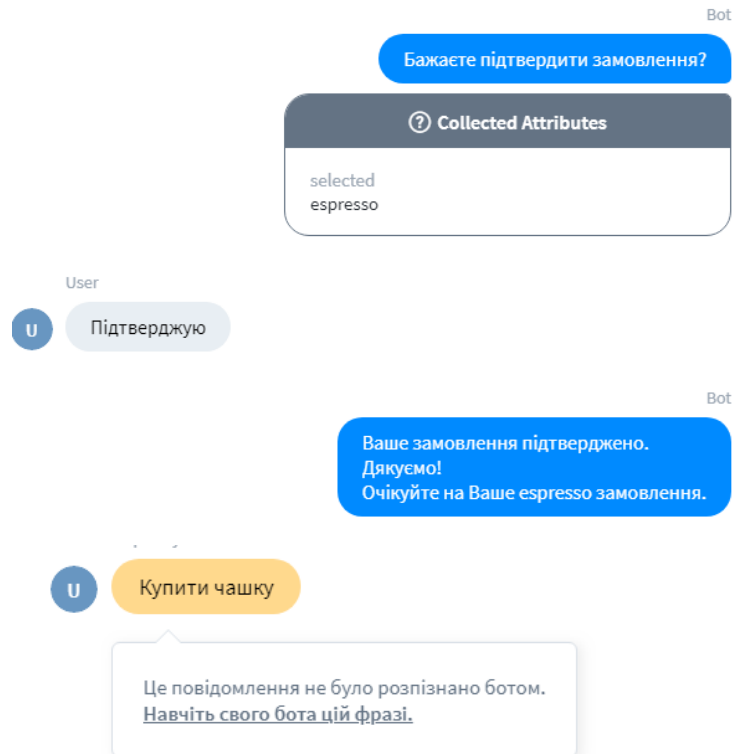


Рисунок 4.6 – Підтвердження замовлення та різка зміна теми

```

{
  "date": "2019-12-15T13:38:40.493Z",
  "requestId": "ba49d4e3-443c-4c85-a16b-5e44e0a376fc",
  "source": "app",
  "resolvedQuery": "Привіт! Я бажую випити кави",
  "confidence": 0.5,
  "score": 0.76,
  "lifespan": 2,
  "incomplete": false,
  "storyId": "5deef5e311add000788535c",
  "interaction": {
    "id": "5deef5e311add000788535e",
    "name": "Замовлення",
    "action": ""
  },
  "parameters": {},
  "fulfillment": [
    {
      "type": "text",
      "delay": 2000,
      "message": "Заісно!\nУ нас є така кава:"
    },
    {
      "type": "cards",
      "delay": 3000,
      "elements": [
        {
          "title": "Еспресо",
          "subtitle": "",
          "imageUrl": "https://images.pexels.com/photos/6347/coffee-cup-workin",
          "buttons": [
            {
              "type": "postback",
              "title": "Обрати",
              "value": "espresso",
              "id": "92a62210-c9be-43e4-b63b-ac95bd756531-ba49d4e3"
            }
          ]
        },
        {
          "title": "Латте",
          "subtitle": "",
          "imageUrl": "https://images.pexels.com/photos/347144/pexels-photo-34",
          "buttons": [
            {
              "type": "postback",
              "title": "Обрати",
              "value": "latte",
              "id": "3434638c-023a-4dc0-9c3b-8e519e60a08f-ba49d4e3"
            }
          ]
        }
      ]
    }
  ]
}

```

Рисунок 4.7 – Розмова в форматі JSON