

Оптимізація голосового пошуку з допомогою IPA

Володимир Грантовський, здобувач¹(ORCID: 0009-0000-9105-241X), Микола Помогаєв, здобувач¹ (ORCID: 0009-0000-1110-9659), Ігор Босенко, д-р філ.¹(ORCID: 0000-0002-9046-4380)

¹Київський національний університет будівництва і архітектури, м. Київ, Україна

АНОТАЦІЯ

У цій роботі розглядається, оптимізація голосового пошуку та розглянуто основні проблеми які виникають при використанні методу ведення голосу у рядок пошуку. Також слід зазначити, що голосовий пошук не просто змінює спосіб введення інформації, але й вимагає переосмислення того, як дані повинні відображатися для користувачів. Короткі та точні відповіді стають ключовими, адже люди звикли отримувати відповіді у форматі розмови. Успішна оптимізація передбачає створення контенту, який легко інтерпретується голосовими асистентами та максимально відповідає запитам користувачів.

Ключові слова: Голосовий пошук, Голосовий помічник, Міжнародний фонетичний алфавіт IPA, нечіткі множини, ступінь належності, Нейронна мережа, глянціотектоніка

1. ВСТУП

У сучасному світі, технології голосового пошуку набирають все більшої популярності. Дана форма взаємодії стала швидким, зручним та інтуїтивним способом отримання інформації або виконання різних завдань. Оптимізація голосового пошуку включає низку методів і технік, що допомагають підвищити шанси як для веб-ресурсу бути знайденим та представленим у результатах пошуку, так і для оперативного та якісного введення інформації з різних мов. У цьому контексті важливо розуміти основні принципи роботи голосових пошукових систем, способи адаптації контенту під голосові запити та роль семантичного пошуку в підвищенні його ефективності.

2. МЕТА РОБОТИ

Метою цієї роботи є дослідження процесів і методів оптимізації голосового пошуку для підвищення його ефективності та точності. Це включає аналіз основних проблем, з якими стикаються користувачі та розробники під час використання голосових систем, і пошук шляхів їх вирішення. Зокрема, дослідження фокусується на покращенні розпізнавання мови, контекстуального розуміння запитів, а також адаптації технологій до багатомовного середовища.

3. ДОСЛІДЖЕННЯ

На сьогоднішній день голосовий пошук активно застосовується багатьма споживачами, особливою популярністю є використання голосового пошуку з мобільних пристроїв. Голосовий пошук — це технологія, яка дозволяє користувачам знаходити інформацію в інтернеті, застосунках або на пристроях за допомогою голосових команд, замість введення тексту [1]. Технологія ґрунтується на розпізнаванні мови та обробці природної мови, що дозволяє системам інтерпретувати сказане і надавати відповідні результати. З плином часу голосовий пошук почали втілювати на основі світових тенденцій на

штучний інтелект, що згодом дало початок поняттю голосового помічника. Голосовий помічник — це програмний додаток, що використовує технології розпізнавання голосу для взаємодії з користувачем.

Одним з популярних помічників, які допомагають знайти інформацію користувачу за допомогою голосу став Google Voice Search, від компанії Google. Працює голосовий помічник за такою схемою: спершу телефон захоплює звуковий сигнал та передає на сервер, на самому, вже сервері проходить аналіз тексту та видача результату, який за допомогою машинного навчання удосконалюється з кожним запитом [2].

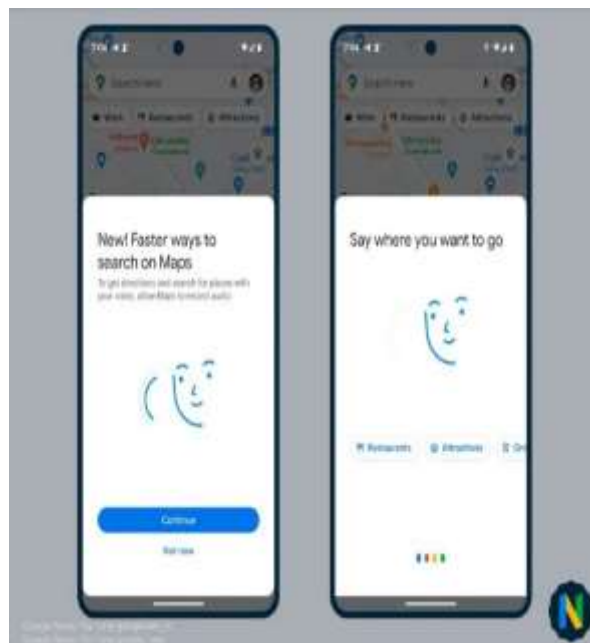


Рисунок 1. Застосування можливостей голосового пошуку на прикладі голосового помічника Google Voice Search

Голосові помічники виконують різноманітні завдання на основі голосових команд та мають такі переваги, як **швидкий доступ до інформації, безконтактна взаємодія, персоналізація, тощо** [3].

4. МЕТОДИ ВИРІШЕННЯ ПРОБЛЕМ ГОЛОСОВОГО ПОШУКУ

4.1. Міжнародний фонетичний алфавіт

Нині значна кількість технологічних компаній інтегрує голосовий пошук у свої додатки та розробки. Найбільша перешкода у впровадженні голосового пошуку яка часто зустрічається ще на стадії впровадження технології це велика кількість звуків, які здатна вимовити людина. Враховуючи що у багатьох мовах світу є безліч різних звуків та вимов літер було узгоджено рішення запровадити Міжнародний фонетичний алфавіт (ІРА), яка дозволяє точно відтворювати кожен звук мови за допомогою стандартизованих символів. Такий підхід істотно підвищує якість роботи систем розпізнавання мовлення, оскільки забезпечує уніфіковане представлення звуків незалежно від діалекту, акценту чи особливостей вимови окремих слів[4].

Використання ІРА сприяє:

- Покращенню точності розпізнавання – системи здатні коректно відрізнити слова зі схожим звучанням.
- Міжмовній інтеграції – алгоритми можуть обробляти декілька мов або діалектів у єдиному форматі, що полегшує глобальне застосування.
- Навчанню моделей – нейронні мережі отримують більш узагальнені й структуровані дані для тренування, що прискорює процес навчання та підвищує ефективність.

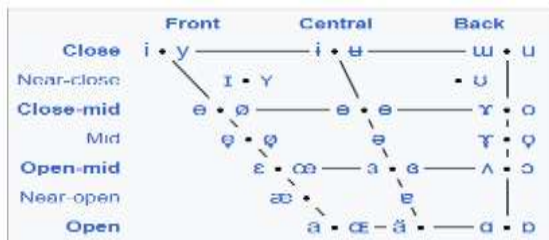


Рисунок 2. Діаграма вимови голосних літер за ІРА

4.2. Застосування нечітких множин

За умови нечітких чи незрозумілих звуків для розвитку розпізнавання слів достатньо залучити математичні методи нечіткої логіки, а саме нечітких множин. У природній мові звуки ніколи не вимовляються абсолютно однаково. Їхні характеристики (частота, інтенсивність, довжина, тембр) коливаються залежно від індивідуальних особливостей мовця, що також впливає на сприйняття програмами голосового пошуку розуміння звуків. Через це межа між звуками є переважно розмитою, один і той самий звук може бути схожим і на [d], і на [t]. В такій ситуації цілком доречно використовувати нечіткі множини.

Звук, який знаходиться між [d] і [t], може мати ступінь належності (міра нечіткої множини) до множини [d] – 0.6 до множини [t] – 0.4. Тобто звук не жорстко класифікується, а оцінюється за рівнем схожості до звуків, передбачених в ІРА. Алгоритми дозволяють чіткіше розпізнавати мовлення: алгоритм не стверджує що «це 100% звук [a]», а використовує нечіткі оцінки: [a] - 0.7, [o] - 0.2, [e] - 0.1. Вибір слова відбувається з урахуванням контексту [5].

4.3. Навчання нейронної мережі

У перспективі очікується ще тісніша інтеграція фонетичних стандартів із технологіями штучного

інтелекту(ШІ). Зокрема, досліджуються методи автоматичного перетворення звукових потоків у фонетичні транскрипції в режимі реального часу, що дозволить значно підвищити швидкість і точність роботи голосових інтерфейсів. Також розглядається можливість створення єдиної багатомовної бази даних фонем, яка слугуватиме основою для універсальних голосових помічників.

На основі сучасного розвитку можливостей штучного інтелекту багато приділяють розвитку нейронних мереж. Нейронна мережа дозволяє здійснювати навчання системи на аудіо даних та розпізнавання мови. Передбачена можливість користування системою через інтерфейс у вигляді чат бота. В процесі розробки були використані технології мови Python та модуль для написання серверів Flask, стандарти машинного навчання нейронних мереж [6]. Також для покращення роботи розпізнавання звуків можна застосувати для навчання нейронної мережі фонетичний алфавіт та методи нечітких множин.

5. ВИСНОВКИ

Використання голосового пошуку набуло популярності завдяки зручності та технологічному прогресу в обробці природної мови. Голосовий пошук надає користувачам можливість шукати інформацію через голосові команди, що робить його швидким і зручним способом взаємодії з пристроями. У таких системах, як Google Assistant, Голосовий пошук продовжує зараз швидко розвиватися, інтегруються різні можливості сприйняття звуків людського мовлення. Розвиток голосового пошуку стає невід'ємною частиною сучасних технологій, що підвищує зручність та ефективність взаємодії з цифровими пристроями.

Список літератури

- [1] Іванов М. Voice Search SEO: оптимізація під голосовий пошук у 2025, URL: https://marketing.ua/blog/article/?id=68&srsltid=AfmBOopMFFa-Z5SEeBjgVHYhCaOmQa9PCfmyVIhctUp-Y1_PipLkew6
- [2] Басюк Т. Просування інтернет-ресурсів з використанням технологій голосового пошуку, URL: <https://science.lpnu.ua/uk/sisn/vsi-vypusky/vypusk-910-2019/prosuвання-internet-resursiv-z-vykorystanniam-tehnologiy>
- [3] Сахарчук Т. Ю. Голосовий асистент для інтегрованих середовищ розроблення програмного забезпечення URL: <https://ela.kpi.ua/server/api/core/bitstreams/0c065c06-3c7a-41b7-8a36-b2328c429b8d/content>
- [4] International Phonetic Alphabet Wiki URL: https://en.wikipedia.org/wiki/International_Phonetic_Alphabet
- [5] Желдак, Т. А. Нечіткі множини в системах управління та прийняття рішень URL: <https://ir.nmu.org.ua/entities/publication/829f0bc0-9a83-4fe1-852a-0ae751eade02>
- [6] Бойко Д Ю Нейромережева система розпізнавання слів в аудіозаписах URL: <https://ela.kpi.ua/server/api/core/bitstreams/65d50146-b1a0-4801-9faf-d8528d55e196/content>