

РОЗВИТОК СИСТЕМ МОВНОЇ АНАЛІТИКИ В ЕКОНОМІЧНІЙ КІБЕРНЕТИЦІ: ІСТОРІЯ, СУЧАСНИЙ СТАН І ПЕРСПЕКТИВИ РОЗВИТКУ

Жуковський Дмитро Миколайович

аспірант кафедри економічної інформатики,

Український державний університет науки і технологій

Транскрибація мовлення, тобто перетворення усної мови в текст, є важливим інструментом в економічній кібернетичі. У контексті економічної інформатики, яка досліджує процеси управління та комунікації в економічних системах, транскрибаційні технології сприяють ефективному збору, обробці та аналізу інформації, що надходить у формі усного мовлення. Сучасні системи перетворення звукових даних у текст відіграють вирішальну роль у бізнес-аналітиці, автоматизації обслуговування клієнтів, створенні контенту, автоматизації контакт-центрів, а також в управлінні знаннями в корпоративному середовищі. Ці технології, зокрема автоматичне розпізнавання мовлення та аналітика мовлення, дозволяють компаніям отримувати цінні інсайти з голосових взаємодій, оптимізувати операційні процеси та підвищувати якість обслуговування клієнтів.

Метою дослідження є аналіз еволюційної траєкторії розвитку систем розпізнавання та аналізу мовлення, а також визначення основних напрямів застосування транскрибаційних технологій та методів обробки природної мови в контексті економічної кібернетики.

На початку 1980-х років розпочався активний процес впровадження систем розпізнавання аудіоданих, що стало можливим завдяки розробці перших комерційних технологій безперервного розпізнавання мовлення. У 1982 році компанія Dragon Systems представила одну з перших систем автоматичного перетворення голосу в текст, яка застосовувала базові принципи статистичного аналізу для підвищення точності розпізнавання мовлення.

Значний технологічний прогрес у 1990-х роках, пов'язаний із вдосконаленням апаратного забезпечення та алгоритмів обробки даних, сприяв створенню більш складних моделей розпізнавання мовлення, зокрема прихованих моделей Маркова, які значно підвищили точність транскрибації завдяки можливості розпізнавати складні акустичні сигнали. Ці моделі стали основою для багатьох комерційних продуктів, включаючи перші версії систем розпізнавання голосу від IBM, Philips і AT&T. Цей період також відзначався значним зростанням інтересу до мовної інженерії та розвитку алгоритмів мовної обробки, що створило фундамент для майбутніх рішень на основі штучного інтелекту.

З початку 2000-х років розвиток потужності комп'ютерних технологій та доступність великих обсягів мовних даних дали поштовх до створення нових поколінь транскрибаційних систем. Одним із перших значних досягнень стала розробка алгоритмів на основі глибоких нейронних мереж *deep neural networks* та рекурентних нейронних мереж *resurgent neural networks*, які могли розпізнавати аудіозаписи значно точніше, ніж традиційні статистичні моделі.

У 2014 році Google запустила першу систему штучного інтелекту транскрибації для комерційного використання, інтегровану в мобільні додатки та хмарні сервіси, такі як Google Speech API. Ця система використовувала великі мовні моделі та векторне представлення слів, що дозволяло досягти високої точності в розпізнаванні контексту.

Подальший розвиток моделей штучного інтелекту для перетворення аудіо в текст став можливим завдяки впровадженню нейронних мереж з трансформерною архітектурою, таких як BERT, GPT і Whisper, які забезпечують вищу точність та контекстну глибину розпізнавання мовлення [1, с. 1]. Whisper, розроблений OpenAI, став важливим проривом у транскрибаційних технологіях, оскільки він базується на масштабованих трансформерних моделях і може працювати з багатьма мовами, включаючи українську, що робить його надзвичайно корисним для міжнародних бізнесів.

Сучасні системи транскрибації використовуються в бізнесі для вирішення широкого спектру завдань, зокрема автоматична транскрибація дзвінків дозволяє створювати детальні записи взаємодії з клієнтами, що використовується для оцінки якості обслуговування, автоматизації роботи операторів та навчання персоналу [2, с. 7].

Автоматичний аналіз розмов із клієнтами допомагає виявляти найефективніші тактики продажів, покращувати скрипти та визначати тригери для успішного завершення угод.

Системи перетворення аудіозаписів у текст значно прискорюють процес обробки матеріалів ділових зустрічей, конференцій та семінарів, забезпечуючи оперативне створення текстових протоколів та документів [3, с. 1095].

Автоматична транскрибація відеоконтенту, подкастів та вебінарів забезпечує кращу індексацію контенту в пошукових системах, що сприяє збільшенню органічного трафіку.

Також у фінансовій сфері моделі для розпізнавання аудіозаписів використовуються для моніторингу розмов із клієнтами з метою дотримання нормативних вимог та запобігання шахрайству [4, с. 17].

У маркетингових дослідженнях транскрибаційні моделі використовуються для аналізу фокус-груп, інтерв'ю та опитувань, що дозволяє отримати точні інсайти про настрої та поведінку споживачів.

У сфері менеджменту персоналом транскрибаційні системи допомагають автоматизувати аналіз співбесід, що підвищує ефективність процесу рекрутингу та покращує якість найму.

Згідно виявлених в результаті проведених досліджень тенденцій, майбутній розвиток транскрибаційних технологій буде зосереджений на таких напрямках: розпізнавання емоцій і намірів в мовному цифровому контенті, інтеграція звітів з аналізу емоцій до комунікаційних систем для більш персоналізованої взаємодії з клієнтами, створення універсальних моделей, здатних розпізнавати різноманітні мови та діалекти, автоматичне створення записів у корпоративних системах управління та використання приватних моделей, що працюють локально для захисту конфіденційної інформації.

Системи розпізнавання звукової інформації пройшли значний еволюційний шлях від використання методів традиційного статистичного аналізу та лінгвістичного моделювання до сучасних технологій штучного інтелекту. Їх розвиток суттєво вплинув на науковий прогрес у галузі економічної кібернетики, створивши нові можливості для автоматизації та підвищення ефективності різноманітних бізнес процесів. Подальші дослідження в цій сфері можуть сприяти розробці ще більш точних, гнучких і безпечних транскрибаційних рішень, що забезпечать вищий рівень персоналізації та покращення клієнтського досвіду.

Список використаних джерел:

1. Chaudhari, K., Dhuri, E., Yadav, V., Vanniyar, D. and Mathur, A., 2024, August. Audio Transcription of Meetings and Action Items using LLM's. In 2024 4th Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-5). IEEE.
2. Getchell, K.M., Carradini, S., Cardon, P.W., Fleischmann, C., Ma, H., Aritz, J. and Stapp, J., 2022. Artificial intelligence in business communication: The changing landscape of research and teaching. *Business and Professional Communication Quarterly*, 85 (1), pp. 7–33
3. Cardon, P.W., Ma, H. and Fleischmann, C., 2023. Recorded business meetings and AI algorithmic tools: Negotiating privacy concerns, psychological safety, and control. *International Journal of Business Communication*, 60 (4), pp. 1095–1122.
4. Crawford Camiciottoli, B., 2019. Multimodality and financial communication: The case of earnings calls. *Ibérica*, 37 (Spring), pp. 17–38.